REMARKS

Under the provisions of 37 CFR 1.136(a), submitted herewith is a petition for a three-month extension of time extending the period for response to the instant Office Action to April 26, 2001. This amendment is therefore timely filed.

The Examiner notes that the priority document, French Patent No. 96 01309, has not been received. In fact, receipt of the priority document was acknowledged by the International Bureau which in turn appears to have forwarded the document to the USPTO as evidenced by the Notification of Acceptance of Application under 35 U.S.C. 371 and 37 CFR 1.494 or 1.495 (copy herewith) wherein receipt of the priority document is acknowledged. If the document has been lost in the PTO, it is respectfully requested that the PTO obtain another copy from the International Bureau under Rule 17.2 of the PCT Regulations.

Claims 1-38 are in the application. In response to the restriction requirement made in the Office Action mailed April 12, 2000 (Paper No. 9), applicants elected the Examiner's Group III, the polypeptide of SEQ ID No. 6 read on by claims 1-6, 25 and 33-36. In a subsequent telephonic restriction requirement claims 6 and 25 were withdrawn from consideration as not having been included in Group III in the restriction requirement of April 12, 2000 and restriction was further required among claims 1-5 and 33-36 as follows:

Group I. Claims 1-5, 33 and 34 drawn to a polypeptide of SEQ ID No. 6

Group II. Claims 35 and 36 drawn to inhibitors of SR-p70 activity.

Applicants elected Group I, claims 1-5, 33 and 34, which election is hereby affirmed.

As to the informalities in the specification noted by the Examiner (items 8-11 of the Office Action), the labeling of the drawings has been amended and the section of the specification originally entitled "LEGEND TO THE FIGURES" has been amended by replacement with a rewritten section entitled "BRIEF DESCRIPTION OF THE DRAWINGS" wherein the title has been changed; reference to the drawings has been amended to reflect the amended numbering of the drawings, and the description of Fig. 9 has been amended to conform to the labeling used in Fig. 9.

Applicants also point out that in the amendment filed January 24, 2000, the description of the drawings was amended to identify each sequence of the drawings.

As to the arrangement of the specification, applicants respectfully submit that all requisite elements of a complete patent application are present and that there is nothing in the law that requires the use of specific headings irrespective of their relevance to the content of a particular application.

Claims 1-5, 33 and 34 are rejected under 35 U.S.C. 101 as lacking a specific utility. The Examiner notes that utilities are disclosed for the elected subject matter, i.e. SEQ ID No. 6 and biologically active fragments thereof as prophylactic, therapeutic and diagnostic agents, in particular, in the field of pathologies linked to the phenomena of apoptosis or of cell transformation, but maintains that the specification does not teach what SEQ ID No. 6 is or what it does, it does not teach a utility for any of the fragments claimed, and does not teach a relationship to any specific disease or establish any involvement in the etiology of any specific diseases. The rejection is traversed and reconsideration thereof is requested. First of all, the specification clearly teaches the identity of SEQ ID No. 6 as human SR-p70 (alternatively known as p73 see WO 99/66946, copy herewith), a polypeptide of 636 amino acids as set forth in Fig. 6A and 6B (specification, p. 17, line 12; drawings pages 10 and 11). As to its utility, the specification points out that SR-p70 (p73) is related to p53 as also noted in Pathol. Int., 2000 Aug.; 50(8):589-93 (copy of abstract herewith). For example, the specification teaches that antibodies to p73 are useful in detecting an abnormal accumulation of p73 proteins in biological samples which makes them useful for detecting cancers or monitoring the progression or remission of pre-existing cancers (specification p 14, line 13, p. 15, line 2). Alternatively, the protein itself can be used to detect auto-antibodies against p73 in patients' sera (specification p. 15, lines 14-17). The accumulation of p73 in tumor cells and the testing for serum antibodies thereto is also documented in the literature, e.g., Br. J. Cancer, 2001 Jan. 5; 84(1):57-63 (copy of abstract herewith). The specification also teaches therapeutic utility in pathologies linked to apoptosis or cell transformation as confirmed in WO 99/66946. Thus, it is submitted that applicants do indeed describe the claimed polypeptides and teach a utility therefor. Withdrawal of the rejection of claims 1-5, 33 and 34 under 35 U.S.C. 101 is therefore requested.

Claims 1-5, 33 and 34 are also rejected under 35 U.S.C. 112, first paragraph. The Examiner maintains that since the claimed invention lacks a utility, one skilled in the art would not know how to use the invention. The rejection is believed overcome in view of the foregoing arguments which clearly establish that applicants do indeed teach a utility for the claimed invention.

The Examiner further urges that the specification does not enable one skilled in the art to make/use the invention commensurate in scope with the claims. The Examiner states that although enabling for SEQ ID No. 6, the specification does not reasonably provide enablement for biologically active sequences derived from SEQ ID No. 6. Applicants disagree. Claim 1 as amended is directed *inter alia* to SEQ ID No. 6 and sequences derived therefrom and having substantially the same biological activity. The specification teaches that derivatives are polypeptides obtained by modification, deletion or addition of a single amino acid or a limited number of amino acids and which are biologically active (specification, p. 3, lines 9-28). The definition of "biologically active" at p. 3, lines 29-37 of the specification substantially mirrors the biological activity described for SEQ ID No. 6. Thus the claimed variants of SEQ ID No. 6, the means of obtaining them and the means of verifying that they have substantially the same biological activity and hence the same utility as SEQ ID No. 6 are well within the skill of the art and would not require undue experimentation. Accordingly, the specification would enable one skilled in the art to practice the claimed invention across its entire scope. Reconsideration and withdrawal of the rejection are therefore requested.

The foregoing arguments also apply to the rejection of claims 33 and 34 under 35 U.S.C. 112, first paragraph. Moreover, one skilled in the art would surely know how to make a pharmaceutical composition containing one of the claimed polypeptides, and since, as noted above, said compounds have substantially the same biological activity, the use thereof in the treatment of cancer is effectively taught. Reconsideration and withdrawal of the rejection are requested.

Claims 1-5, 33 and 34 are also rejected under 35 U.S.C. 112, first paragraph on the grounds that the specification fails to describe the claimed invention. Substantially for the reasons given above, in particular with regard to the description of sequences derived from SEQ ID No. 6 at page 3, lines 9-37, it is submitted that the

specification does convey to one skilled in the art applicant's claimed invention. Reconsideration and withdrawal of the rejections are requested.

Claims 4, 33 and 34 are rejected under 35 U.S.C. 112, second paragraph as being indefinite. The rejection is believed fully met by the amendments of said claims.

Claims 1-4, 33 and 34 are rejected under 35 U.S.C. 102(b) as being anticipated by Dequiedt et al., (DNA SEQ. 11995, 5:255-259). The rejection is traversed and reconsideration thereof is requested. As described in applicants' specification (p. 3, lines 9-28, variant polypeptides are those obtained by modification, addition or deletion of a single amino acid, or of a limited number of amino acids as well as any isoform sequence and which, as specified in the claims, have substantially the same biological activity as SEQ ID No. 6. The cited reference shows a comparison of the prior art database polypeptide with a p73 partial sequence of some 350 amino acids. This is hardly a variant involving a single or a limited number of the 636 amino acids of p73. Moreover, even taking the entire sequence shown there are nearly as many mismatches as there are matches. Clearly then, the reference simply does not teach applicants' polypeptides and accordingly is incompetent to anticipate the instant claims. Reconsideration and withdrawal of the rejections under 102(b) is requested.

Claims 1 and 5 are rejected under 35 U.S.C. 103 as being unpatentable over Dequiedt et al. in view of U.S. Patent 5,532,348. Reconsideration and withdrawal are requested.

It is submitted that since, as pointed out hereinabove, the primary reference fails to teach the claimed polypeptides the secondary reference relating to a p53 fusion protein adds nothing to the primary reference and hence the combined references would not have suggested a fusion protein of the instantly claimed polypeptides.

Attached hereto is a marked-up version of the changes made to the specification and claims by the current amendment. The attachment is entitled "Version With Markings To Show Changes Made".

There being no remaining issues, this application is believed in condition for favorable reconsideration and such action is earnestly solicited.

Respectfully submitted,

Date: 4/29/01

Paul E. Dupont
Reg. No. 27,438

Address:
Patent Department
Sanofi-Synthelabo Inc.
9 Great Valley Parkway
Malvern, PA 19355
Telephone No. (610) 889-6338
Facsimile: (610) 889-8799

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In The Specification:

The section entitled "LEGEND TO THE FIGURES" at page 16, line 34 to page 19, line 32 has been deleted and replaced by the following section:

[LEGEND TO THE FIGURES] BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1A and 1B [Figure 1]:	Nucleic acid comparison of monkey SR-p70a cDNA (corresponding to nucleotides 1-1599 of SEQ ID No. 1) with the nucleic acid sequence of monkey p53 cDNA (SEQ ID No. 43).
[Figure] Fig. 2:	Protein comparison of monkey SR-p70a amino acids 1-450 of SEQ ID No. 1 with monkey p53 protein (SEQ ID No. 44) (sw: p53-cerae).
Fig. 3A-C [Figure 3]:	Comparison of the nucleic acid sequence of monkey SR-p70a and b cDNA (corresponding, respectively, to SEQ ID No. 1 and SEQ ID No. 3).
Fig. 4A and 4B [Figure 4]:	Nucleic acid sequence (SEQ ID No. 1) and deduced protein sequence (SEQ ID No. 2) of monkey SR-p70a.
[Figure] Fig. 5:	Partial nucleic acid sequence (SEQ ID No. 3) and complete deduced protein sequence (SEQ ID No. 4) of monkey SR-p70b.
Fig. 6A and 6B [Figure 6]:	Partial nucleic acid sequence (SEQ ID No. 5) and deduced complete protein sequence (SEQ ID No. 6) of human SR-p70a.
[Figure] Fig. 7:	Partial nucleic acid sequence (SEQ ID No. 7) and complete deduced protein sequence (SEQ ID No. 8) of mouse SR-p70c.
[Figure] Fig. 8:	Partial nucleic acid sequence (SEQ ID No. 9) and partially deduced protein sequence (SEQ ID No. 10) of mouse SR-p70a.
Fig. 9A and 9B [Figure 9]:	Multialignment of the proteins deduced from monkey (SR-p70a-cos3 and SR-p70b-cos3) [(a and b)] (SEQ ID

No. 2 and SEQ ID No. 4, respectively), human [(a)] (SR-p70-ht29) and mouse [(a and c)] (SR-p70c-att20 and sr-p70a-att20) (SEQ ID No. 10 and SEQ ID No. 8, respectively), SR-p70 cDNAs.

[Figure] Fig. 10a: Immunoblot of the SR-p70 protein.

[Figure] Fig. 10b: Detection of the endogenous SR-p70 protein.

[Figure] Fig. 11: Chromosomal localization of the human SR-p70 gene. The signal appears on chromosome 1, in the p36 region.

[Figure] Fig. 12: Genomic structure of the SR-p70 gene and comparison with that of the p53 gene. The human protein sequences of SR-p70a (SEQ ID No. 6) (upper line of the alignment) and of p53 (SEQ ID No. 45) (lower line) are divided up into peptides on the basis of the respective exons from which they are encoded. The figures beside the arrows correspond to the numbering of the corresponding exons.

[Figure] Fig. 13: Human genomic sequence of SR-p70 from the 3' end of intron 1 to the 5' end of exon 3 (SEQ ID No. 46). The introns are boxed. At positions 123 and 133, two variable nucleic acid positions are localized (G □ A at 123 and C □ T at 133). The restriction sites for the enzyme StyI are underlined (position 130 in the case where a T is present instead of a C at position 133, position 542 and position 610). The arrows indicate the positions of the nucleic acid primers used in Example XI.

[Figure] Fig. 14: Nucleic acid comparison of the 5' region of the human cDNAs of SR-p70d (SEQ ID No. 12) and of SR-p70a (SEQ ID No. 5).

Fig. 15A-J [Figure 15]: Multialignment of the nucleic acid sequences corresponding to human SR-p70a, b, d, e, and f (SEQ ID No. 5, SEQ ID No. 18, SEQ ID No. 12, SEQ ID No. 14 and SEQ ID No. 16, respectively).

Fig. 16A-C [Figure 16]: Multialignment of the proteins deduced from human SR-p70 (a, b, d, e and f) (SEQ ID No. 6, SEQ ID No. 19, SEQ ID No. 13, SEQ ID No. 15 and SEQ ID No. 17, respectively), cDNA's.

[Figure] Fig. 17: Partial nucleic acid sequence (SEQ ID No. 5) and partial deduced protein sequence (SEQ ID No. 6) of human SR-p70a. The two bases in bold characters correspond to two variable positions (see Figure 6). This sequence

possesses a more complete non-coding 5' region than the one presented in Figure 6.

[Figure] Fig. 18: Analysis of the SR-p70a transcripts after PCR amplification.
lane M: 1 kb ladder (GIBCO-BRL) molecular weight markers
lane 1: line HT29
lane 3: line SK-N-AS
lane 5: line UMR-32
lane 7: line U-373 MG
lane 9: line SW 480
lane 11: line CHP 212
lane 13: line SK-N-MC

lanes 2, 4, 6, 8, 10, 12, 14: negative controls corresponding to lanes 1, 3, 5, 7, 9, 11 and 13, respectively (absence of inverse transcriptase in the RT-PCR reaction).

Fig. 19A and 19B [Figure 19]: A: Analysis by agarose gel electrophoresis of genomic fragments amplified by PCR (from the 3' end of intron 1 to the 5' end of exon 3). The numbering of the lanes corresponds to the numbering of the control population. Lane M: molecular weight markers (1 kb ladder).

B: Analysis identical to that of part A, after digestion of the same samples with the restriction enzyme StyI.

[Figure] Fig. 20: Diagrammatic representation with a partial restriction map of the plasmid pCDNA3 containing human SR-p70a.

In The Claims:

Claims 1, 4 and 33 have been amended as follows:

1.(Twice amended) A purified polypeptide, comprising an amino acid sequence selected from the group consisting of:

a) sequence SEQ ID No. 2;

b) sequence SEQ ID No. 4;

c) sequence SEQ ID No. 6;

d) sequence SEQ ID No. 8;

e) sequence SEQ ID No. 10;

f) sequence SEQ ID No. 13;

g) sequence SEQ ID No. 15;

h) sequence SEQ ID No. 17;

i) sequence SEQ ID No. 19; and

j) any [biologically active] sequence derived from SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19 and having substantially the same biological activity.

4. (Twice amended) A polypeptide according to Claim 1, which is produced from an alternative splicing of messenger RNA of a gene coding for said polypeptide.

33. (Twice amended) A pharmaceutical composition for the treatment of pathologies linked to apoptosis or cell transformation comprising an effective amount of the polypeptide according to Claim 1.

1: Br J Cancer 2001 Jan 5;84(1):57-63 Related Articles, Books, LinkOut

IDEAL

Detection of p73 antibodies in patients with various types of cancer: immunological characterization.

Tominaga O, Unsal K, Zalcman G, Soussi T.

Unite de genotoxicologie des tumeurs, Institut Curie, 26 rue d'Ulm, Paris, 75005.

p53 antibodies have been found in the sera of patients with various types of cancer. The presence of these antibodies is generally associated with p53 accumulation in the tumour that is believed to trigger this humoral response. The recent discovery of 2 new members of the p53 family, p73 and p63, led us to study the specificity of this immune response towards the 3 proteins. Serum samples from 148 patients with various types of cancer were tested for antibodies against p73 and p63 using immunoprecipitation. 72 patients were previously shown to have p53 antibodies whereas 76 were negative. The control group consisted of 50 blood donors. p73 were detected in 22/148 (14.9%) of the cancer patients (11/72 in the group with p53-antibodies and 11/76 in the negative group). Only two sera from the control (4%) were positive. p63 antibodies were detected in only 4/148 (2.7%) of the cancer patients. Epitope mappings were performed and demonstrate that p73 antibodies are directed toward the central region of the p73 protein whereas p53 antibodies react predominantly toward the amino- and the carboxy-terminus of p53. Our results indicate that there is a specific immune response toward the p73 protein in cancer patients, a finding supported by an increasing number of publications describing p73 accumulation in tumoral cells.

PMID: 11139314 [PubMed - indexed for MEDLINE]

Display Abstract Save Text Order Add to Clipboard

1: Pathol Int 2000 Aug;50(8):589-93 Related Articles, Books, LinkOut

online

p73: structure and function.

Ichimiya S, Nakagawara A, Sakuma Y, Kimura S, Ikeda T, Satoh M, Takahashi N, Sato N, Mori M.

Department of Pathology, Sapporo Medical University School of Medicine, Sapporo, Japan. ichimiya@sapmed.ac.jp

Alteration of the p53 tumor suppressor gene is a common, if not general, observation in human malignant tumors. p73 Is a novel member of the p53 family at chromosome 1p36.3, at which locus frequent defects are seen in many tumors including neuroblastoma. Besides structural similarities, the fact that p73 functions in the regulation of the cell cycle and apoptosis promotes the expansion of the research field concerning p53-associated tumor progression. In this paper, we review the structure and function of p73 as well as the mutational status in various human tumors. In addition, possibilities for new therapeutic applications with p73 for cancer cell control are discussed.

Publication Types:
- Review
- Review, tutorial

PMID: 10972855 [PubMed - indexed for MEDLINE]

Display Abstract Save Text Order Add to Clipboard

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A61K 38/00, 48/00, C12N 15/85

A1

(11) International Publication Number: WO 99/66946

(43) International Publication Date: 29 December 1999 (29.12.99)

(21) International Application Number: PCT/US99/14057

(22) International Filing Date: 23 June 1999 (23.06.99)

(30) Priority Data:
60/090,526 24 June 1998 (24.06.98) US

(71) Applicant *(for all designated States except US)*: TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA [US/US]; Suite 300, 3700 Market Street, Philadelphia, PA 19104 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: EL-DEIRY, Wafik, S. [US/US]; Apartment P113, 1500 Locust Street, Philadelphia, PA 19102 (US).

(74) Agents: REED, Janet, E. et al.; Dann Dorfman Herrell and Skillman, Suite 720, 1601 Market Street, Philadelphia, PA 19103 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: COMPOSITIONS AND METHODS FOR INDUCING APOPTOSIS IN E6-EXPRESSING CELLS

(57) Abstract

Methods, pharmaceutical compositions and kits are provided for inducing programmed cell death in cells expressing the E6 oncogene. The methods and compositions are particularly suited for treatment of cancers involving infections with E6-expressing virus, such as human papilloma virus (HPV). The methods and compositions utilize the p53 homolog, p73. Unlike p53, p73 is not targeted by the E6 oncoprotein for ubiquitin-mediated degradation, and so provides a viable alternative to p53 therapy for treatment of E6-expressing cancers.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

COMPOSITIONS AND METHODS FOR INDUCING APOPTOSIS IN E6-EXPRESSING CELLS

This application claims priority to U.S. Provisional Application 60/090,526, filed June 24, 1998, the entirety of which is incorporated by reference herein.

FIELD OF THE INVENTION

This invention relates to the field of methods of treatment of cancer. In particular, this invention provides a method of treatment of cancers associated with human papillomavirus infection or other tumors in which the E6 oncogene is expressed, and a pharmaceutical preparation and kit to practice the method.

BACKGROUND OF THE INVENTION

Various scientific and scholarly articles and patents are referred to in brackets throughout the specification. These articles and patents are incorporated by reference herein to describe the state of the art to which this invention pertains.

Infection with human papillomavirus (HPV) is a major risk factor for the development of squamous cell carcinoma of the cervix. The E6-oncoprotein encoded by HPV has been shown to target the tumor suppressor protein p53 for degradation *via* ubiquitin conjugation and subsequent proteolysis (Scheffner et al., 1990, Cell 63: 1129-1136). HPV-E6-expressing cancer cells are resistant to the tumor suppressive effects of exogenous wild-type

p53 delivered by an adenovirus (Ad) vector (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229).

Several approaches have been proposed to control the growth of HPV E6-expressing cancer cells. These include the use of p21-expressing adenovirus to bypass the p53-degradation step (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229), the use of bovine papillomavirus E2 gene to reactivate endogenous p53 (Hwang et al., 1996, Oncogene 12:795-803), the use of hypoxic conditions to suppress p53 degradation (Kim et al., 1997, Cancer Res. 57:4200-4204), the use of alternatively spliced E6 to compete with normally spliced E6 (Pim et al., 1997, Oncogene 15: 257-264), the use of antisense strategies to lower E6 expression (Hamada et al., 1996, Gyn. Onc. 63:219-227; Beer-Romero et al., 1997, Oncogene 14: 595-602), and the generation of p53 mutants resistant to ubiquitin-directed degradation (Crook et al., 1996, Virology, 217:285-292). In the last mentioned approach, it was found that, although lysine mutants of the C-terminus of p53 did resist E6-mediated degradation *in vitro*, the effect was not observed in intact cells, where the lysine mutant was efficiently targeted for degradation (Crook et al., 1996, Virology, 217:285-292). Some tumor-derived mutants of p53 may also be resistant to E6-dependent proteolysis *in vitro* (Medcalf and Milner, 1993, Oncogene 8:2847-2851). p21-expressing adenovirus (Ad-p21) inhibits the growth of E6-over-expressing cells, although the primary effect of p21 over-expression is a growth arrest associated with a large cell phenotype and little, if any, apoptosis (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229; Meng et al., 1998, Clin. Cancer Res. 4: 251-259). Thus, none of the aforementioned approaches has been

particularly successful in controlling the growth of E6-over-expressing cells by induction of programmed cell death.

Alternative strategies for the suppression of growth of E6-expressing cancer cells are of great utility. Such alternative strategies would ideally induce apoptosis of the E6-over-expressing cells, as well as inhibit cell proliferation.

SUMMARY OF THE INVENTION

Therapy based on the p53 tumor suppressor is unavailable for cancers associated with expression of the E6 oncogene because the E6 protein targets p53 for degradation by ubiquitin-mediated proteolysis. It has been discovered in accordance with the present invention that the p53 homolog, p73, is not targeted for degradation by E6 and, moreover, is a potent inhibitor of cancer colony growth and inducer of apoptosis, even in cells that over-express E6. Thus, p73 is a superior tumor suppressor protein for treatment of cancers in which the E6 oncogene is expressed, such as those associated with HPV infection.

According to one aspect of the present invention, method is provided for inducing apoptosis in an E6-expressing cell. The method comprises administering to the cell an amount of p73 protein effective to induce the apoptosis. In one embodiment, the p53 protein is administered as a DNA construct comprising an expressible sequence that encodes the protein. Preferably, the DNA construct is operably inserted into a viral vector for transforming cells.

The method is typically utilized for arresting growth of cancerous cells, particularly cancers

associated with infection with E6-expressing viruses, such as HPV. In one embodiment, the cell is a cultured cell In another embodiment, the cell is obtained from the body of a living organism, the administering is performed *ex vivo*, and the cell is returned to the living organism. In still another embodiment, the cell is disposed within a living organism and the administering is performed *in vivo*.

The p73 protein utilized in the method is preferred to be p73α or p73β, most preferably the latter. In a preferred embodiment, the protein comprises a sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:2. If a DNA construct is used, the DNA construct preferably comprises more than 50 nucleotides of SEQ ID NO:3.

According to another aspect of the invention, an apoptotic, E6-expressing transgenic cell is provided, which comprises a heterologous, expressible DNA construct encoding p73. In one embodiment, the cell is obtained from a cultured cell line. In another embodiment, the cell is a primary cell obtained from a living organism. In yet another embodiment, it is disposed within a living organism.

According to another aspect of the invention, a pharmaceutical preparation for treatment of cancers associated with E6 over-expression is provided. In one embodiment, the pharmaceutical preparation comprises a p73 protein associated with a delivery vehicle for delivering proteins to cancer cells. In another embodiment, the preparation comprises an expressible DNA construct encoding p73, associated with a delivery vehicle for delivering DNA to cancer cells. The

pharmaceutical preparation also may comprise at least one additional active ingredient for treatment of cancer.

According to another aspect of the invention, a kit is provided that contains the pharmaceutical preparation and other optional components. For instance, in a preferred embodiment, the kit may include a second pharmaceutical agent useful for treating cancer.

Other features and advantages of the present invention will be better understood by reference to the drawings, detailed description and examples that follow.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Ad-E6 infection leads to degradation of both wild-type and mutant p53 in human cancer cells. The human brain (U373, H80; lanes 1-4), breast (SKBr3; lanes 5,6), lung (H460; lanes 7,8), or colon (HCT116, SW480; lanes 9-12) cancer cell lines were infected using Ad-LacZ or Ad-E6 (as indicated). Immunoblotting for p53 expression (upper panels) or pRb expression (lower panels) was carried out as described in Example 1. pRb expression is presented to document equivalent loading between lysates derived from Ad-LacZ and Ad-E6 infected cells. For cell lines that express mutant p53, the following mutations have been previously reported: U373 cell line: R273H (Kaghad et al., 1997, Cell 90: 809-819); SW480 cell line: R273H, P309S (Kaghad et al., 1997, Cell 90: 809-819); SKBr3 cell line: R175H (Kovach et al., 1991, J. Natl. Cancer Inst., 83:1004-1009); H80 cell line (also known as U-373 MG): R273H (Gomez-Manzano et al., 1996, Cancer Res. 56:694-699).

Figure 2. p73, unlike p53, is not specifically targeted for degradation in Ad-E6 infected cancer cells. SW480 cells were transfected by p73α (lanes 1,2), p73αm

(lanes 3,4), p73β (lanes 5,6), or p73βm (lanes 7,8). At six hours following transfection, cells were infected by either Ad-LacZ or Ad-E6 (as indicated). At 20 hrs. following infection, expression of p73α (upper left) or p73β (upper right) was detected by immunoblotting using anti-HA antibody and for p53 (lower panels) expression using anti-p53 antibody, as described in Example 1. The band just above p73α is a non-specific anti-HA cross-reactive band.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

Various terms relating to the biological molecules of the present invention are used hereinabove and also throughout the specifications and claims.

With reference to nucleic acid molecules, the term "isolated nucleic acid" is sometimes used. This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous (in the 5' and 3' directions) in the naturally occurring genome of the organism from which it was derived. For example, the "isolated nucleic acid" may comprise a DNA molecule inserted into a vector, such as a plasmid or virus vector, or integrated into the genomic DNA of a procaryote or eucaryote. An "isolated nucleic acid molecule" may also comprise a cDNA molecule.

With respect to RNA molecules, the term "isolated nucleic acid" primarily refers to an RNA molecule encoded by an isolated DNA molecule as defined above. Alternatively, the term may refer to an RNA molecule that has been sufficiently separated from RNA molecules with which it would be associated in its natural state (i.e., in cells or tissues), such that it

exists in a "substantially pure" form (the term "substantially pure" is defined below).

With respect to proteins or polypeptides, the term "isolated protein (or polypeptide)" or "isolated and purified protein (or polypeptide)" is sometimes used herein. This term refers primarily to a protein produced by expression of an isolated nucleic acid molecule of the invention. Alternatively, this term may refer to a protein which has been sufficiently separated from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form.

The term "substantially pure" refers to a preparation comprising at least 50-60% by weight the compound of interest (e.g., nucleic acid, oligonucleotide, protein, etc.). More preferably, the preparation comprises at least 75% by weight, and most preferably 90-99% by weight, the compound of interest. Purity is measured by methods appropriate for the compound of interest (e.g. chromatographic methods, agarose or polyacrylamide gel electrophoresis, HPLC analysis, and the like).

Nucleic acid sequences and amino acid sequences can be compared using computer programs that align the similar sequences of the nucleic or amino acids thus define the differences. For purposes of this invention, the GCG Wisconsin Package version 9.1, available from the Genetics Computer Group in Madison, Wisconsin, and the default parameters used (gap creation penalty=12, gap extension penalty=4) by that program are the parameters intended to be used herein to compare sequence identity and similarity. Alternatively, standard BLAST query parameters, utilized by public databases such as GenBank, are utilized herein.

The term "substantially the same" refers to nucleic acid or amino acid sequences having sequence variation that do not materially affect the nature of the protein (i.e. the structure, thermostability characteristics and/or biological activity of the protein). With particular reference to nucleic acid sequences, the term "substantially the same" is intended to refer to the coding region and to conserved sequences governing expression, and refers primarily to degenerate codons encoding the same amino acid, or alternate codons encoding conservative substitute amino acids in the encoded polypeptide. With reference to amino acid sequences, the term "substantially the same" refers generally to conservative substitutions and/or variations in regions of the polypeptide not involved in determination of structure or function.

The terms "percent identical" and "percent similar" are also used herein in comparisons among amino acid and nucleic acid sequences. When referring to amino acid sequences, "percent identical" refers to the percent of the amino acids of the subject amino acid sequence that have been matched to identical amino acids in the compared amino acid sequence by a sequence analysis program. "Percent similar" refers to the percent of the amino acids of the subject amino acid sequence that have been matched to identical or conserved amino acids. Conserved amino acids are those which differ in structure but are similar in physical properties such that the exchange of one for another would not appreciably change the tertiary structure of the resulting protein. Conservative substitutions are defined in Taylor (1986, J. Theor. Biol. 119:205). When referring to nucleic acid molecules, "percent identical" refers to the percent of

the nucleotides of the subject nucleic acid sequence that have been matched to identical nucleotides by a sequence analysis program.

Transcriptional and translational control sequences, sometimes referred to herein as "expression control" sequences or elements, or "expression regulating" sequences or elements, are DNA regulatory elements such as promoters, enhancers, ribosome binding sites, polyadenylation signals, terminators, and the like, that provide for the expression of a coding sequence in a host cell. The term "expression" is intended to include transcription of DNA and translation of the mRNA transcript.

The terms "promoter", "promoter region" or "promoter sequence" refer generally to transcriptional regulatory regions of a gene, which may be found at the 5' or 3' side of the coding region, or within the coding region, or within introns. Typically, a promoter is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. The typical 5' promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence is a transcription initiation site (conveniently defined by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

The term "selectable marker gene" refers to a gene encoding a product that, when expressed, confers a

selectable phenotype such as antibiotic resistance on a transformed cell.

The term "operably linked" means that the regulatory sequences necessary for expression of a particular coding sequence are placed in the DNA molecule in the appropriate positions relative to the coding sequence so as to enable expression of the coding sequence. This same definition is sometimes applied to the arrangement of transcription units and other regulatory elements (e.g., enhancers or translation regulatory sequences) in an expression vector.

A "vector" is a replicon, such as plasmid, phage, cosmid, or virus to which another nucleic acid segment may be operably inserted so as to bring about the replication or expression of the segment.

The term "nucleic acid construct" or "DNA construct" is sometimes used to refer to a coding sequence or sequences operably linked to appropriate regulatory sequences and inserted into a vector for transforming a cell. This term may be used interchangeably with the term "transforming DNA". Such a nucleic acid construct may contain a coding sequence for a gene product of interest, along with a selectable marker gene and/or a reporter gene.

A "heterologous" region of a nucleic acid construct is an identifiable segment (or segments) of the nucleic acid molecule within a larger molecule that is not found in association with the larger molecule in nature. Thus, when the heterologous region encodes a mammalian gene, the gene will usually be flanked by DNA that does not flank the mammalian genomic DNA in the genome of the source organism. In another example, coding sequence is a construct where the coding sequence itself

is not found in nature (e.g., a cDNA where the genomic coding sequence contains introns, or synthetic sequences having codons different than the native gene). Allelic variations or naturally-occurring mutational events do not give rise to a heterologous region of DNA as defined herein.

A cell has been "transformed" or "transfected" by exogenous or heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. For example, the transforming DNA may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA. A "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

"Killing", "programmed cell death" and "apoptosis" are used interchangeably in this text to describe a series of cellular events that culminates in the death of the target cell. Apoptosis is a characteristic morphological change in which the cell and its nucleus shrink, condense and fragment. Frequently accompanying this morphological change are the activation of intracellular proteases and nucleases that lead to, for example, cell nucleus involution and nuclear DNA fragmentation.

II. Description

Human papillomavirus (HPV) is the major cause of cervical cancer worldwide. HPV-E6 protein targets the p53 tumor suppressor protein for degradation by ubiquitin-mediated proteolysis, making such cancers resistant to p53-mediated therapy.

In accordance with the present invention, two discoveries have been made that have significant implications and suggest novel strategies for cancer therapy. The first discovery is that HPV-E6 targets both endogenous wild-type and mutant p53 for degradation (Fig. 1). Possibly because p53 mutations are rare in cervical cancer (Busby-Earle et al., 1994, Br. J. Cancer 69: 732-737) the hypothesis that HPV-E6 could target endogenous mutant p53 for degradation has not been previously directly tested. While several studies have reported low levels of p53 expression and an inverse correlation between the presence of HPV and p53 expression (Scheffner et al., 1991, Proc. Natl. Acad. Sci. USA 88: 5523-5527; Srivastava et al., 1992, Carcinogenesis 13: 1273-1275; Baret al., 1996, Eur. J. Gyn. Onc. 17: 283-285; Hachisuga et al., 1996, Pathology 28: 28-31), there is apparently no such correlation between p53 mutation and HPV (Busby-Earle et al., 1994, Br. J. Cancer 69: 732-737; Kim and Kim, 1995, Yonsei Med. J. 36:412-425). While the discovery that HPV-E6 also targets mutant p53 for degradation provides no insight into how the rare p53 mutations may contribute to HPV-associated cervical cancer, it is consistent with the known inverse correlation between p53 expression and the presence of HPV in high risk cervical cancer.

The second and more significant discovery is that the p53 homolog, p73, is not targeted for

degradation by the E6 oncoprotein. Furthermore, as described in greater detail below and in Example 1, p73 is a potent inducer of apoptosis and is an effective inhibitor of cancer cell growth. For such HPV E6-expressing cancers where p53 is degraded and fails to control growth, p73 is an excellent substitute for p53 in gene replacement because of its resistance to E6-mediated proteolysis. Other differences of effect of viral oncoproteins have been noted (Marin et al., 1998, Mol. Cell. Biol. 18:6316-6324; Steengenga et al., 1999, Mol. Cell. Biol. 19:3885-3894; Dobbelstein and Roth, 1998, J. Gen Virol. 79:3079-3083; Roth et al., 1998, J. Virol. 72:8510-8516; Reichelt et al., 1999, Arch. Virol. 144:621-626). It is noteworthy that, even though p73 has the potential to interact with p53 in a yeast two-hybrid analysis (Kaghad et al., 1997, Cell 90: 809-819), the expressed p73 is not subject to E6-dependent proteolysis under conditions where high levels of endogenous mutant p53 are degraded (Fig. 2).

Provided with this invention are methods, pharmaceutical preparations and kits that utilize p73 for arresting the growth of E6-expressing cells, particularly HPV-infected cancer cells. The treatment of the target cells may be *in vivo*, within the patient; or *ex vivo*, removed from the patient, treated, and reintroduced into the patient. It is contemplated that the methods, pharmaceutical preparations and kits of the invention can be used alone or in conjunction with chemotherapy or radiation therapy to treat cancers *in vivo*. Additionally, the methods, pharmaceutical preparations and kit of the invention can be used for experimental purposes *in vitro* with standard cell cultures.

As mentioned, the treatment of cancers associated with the over-expression of E6 protein is of particular interest. Several circumstances may result in mammalian cells that over-express E6 protein. Commonly, this nature of cell results from an infection with human papillomavirus (HPV) wherein the E6-oncogene encoded by the virus is expressed in the cell. HPV infection is well-known to result in cancers of the uterine cervix. In addition to anogential cancer, HPV infection may also result in esophageal squamous cell cancer, laryngeal papilloma, bronchiolo-alveolar carcinoma, penile carcinoma and bladder carcinoma, among others. Additionally, E6-over-expression may also result from a mutation in the mammalian cell genome such that the endogenous E6 gene is over-expressed. All mammalian cells that over-express the E6 protein, regardless of the origin of the phenotype, are contemplated for treatment with the method of the invention.

The following description set forth the general procedures involved in practicing the present invention. To the extent that specific materials are mentioned, it is merely for purposes of illustration and is not intended to limit the invention. Unless otherwise specified, general cloning procedures, such as those set forth in Sambrook et al., Molecular Cloning, Cold Spring Harbor Laboratory (1989) (hereinafter "Sambrook et al.") or Ausubel et al. (eds) Current Protocols in Molecular Biology, John Wiley & Sons (1999) (hereinafter "Ausubel et al.") are used.

Any p73 variant and the nucleic acid sequence encoding it are considered suitable for use in the present invention. In this regard, it should be noted that the two major splice variants of p73, p73α and p73β,

both have been found resistant to E6-mediated degradation (see Example 1), though p73β appears to be somewhat more effective in this regard and is preferred for the practice of the present invention.

The amino acid sequence of p73 protein on which to base the nucleic acid construct is ideally from the gene that is endogenous to the species which is being treated. In a preferred embodiments, *Homo sapiens* is being treated and the nucleic acid construct encodes SEQ ID NO:1 or SEQ ID NO:2. In a most preferred embodiment, the nucleic acid sequence is SEQ ID NO:3. Other variants of p73 protein also exist in *Homo sapiens* and the sequences of these variants are also contemplated for use with the invention (DeLaurenzi et al., 1999, Cell Death Differ. 6:389-390 incorporated by reference herein; Genbank Accession No. Y11416 incorporated by reference herein).

The availability of amino acid sequence information, such as the full length sequence in SEQ ID NO:1 and SEQ ID NO:2 enables the preparation of a synthetic gene that can be used to synthesize the *Homo sapiens* p73 protein in standard *in vivo* expression systems or to make viral vectors expressing the p73 protein. The sequence encoding *Homo sapiens* p73 from isolated native nucleic acid molecules such as SEQ ID NO:3 can be utilized. The amino acid and nucleic acid sequences found in Genbank Accession Nos. AF138873, Y11419 and AF043641 can be used to prepare the p73 protein endogenous to *Mus musculus*, *Chlorocebus aethiops* and *Barbus barbus*, respectively. Alternately, an isolated nucleic acid that encodes the amino acid sequence of the invention can be prepared by oligonucleotide synthesis. Codon usage tables can be

used to design a synthetic sequence that encodes the protein of the invention. In a preferred embodiment, the codon usage table has been derived from the organism in which the synthetic nucleic acid will be expressed. For example, the codon usage for *E. coli* would be used to design an expression DNA construct to produce the *Homo sapiens* p73 in *E. coli*.

Synthetic oligonucleotides may be prepared by the phosphoramadite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant oligonucleotide may be purified according to methods known in the art, such as high performance liquid chromatography (HPLC).

Nucleic acid molecules encoding p73 also may be isolated from appropriate species using methods well known in the art. Native nucleic acid sequences may be isolated by screening mammalian or other cDNA or genomic libraries with oligonucleotides preferably designed to match the *Homo sapiens* coding sequence of p73 (SEQ ID NO:3). Several other p73 amino acid sequences are now known: *Mus musculus*, Genbank Accession No. AF138873; *Chlorocebus aethiops* (Green Monkey), Genbank Accession No. Y11419; and *Barbus barbus*, Genbank Accession No. AF043641; each of these sequences is incorporated by reference herein. Oligonucleotides designed to match any of these sequences or to match regions of high homology between these sequences may also be used to screen for mammalian p73-encoding nucleotides. In positions of degeneracy where more than one nucleic acid residue could be used to encode the appropriate amino acid residue, all the appropriate nucleic acids residues may be incorporated to create a mixed oligonucleotide population, or a neutral base such as inosine may be

used. The strategy of oligonucleotide design is well known in the art (see also Sambrook et al., Molecular Cloning, 1989, Cold Spring Harbor Press, Cold Spring Harbor NY). Alternatively, PCR (polymerase chain reaction) primers may be designed by the above method to match a known coding sequence of p73, and these primers used to amplify the native nucleic acids from isolated mammalian cDNA or genomic DNA.

Nucleic acids having the appropriate sequence homology with a *Homo sapiens* p73 synthetic nucleic acid molecule may be identified by using hybridization and washing conditions of appropriate stringency. One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology (Sambrook et al., 1989, *supra*):

$$T_m = 81.5°C + 16.6\text{Log}[Na+] + 0.41(\%\ G+C) - 0.63\ (\%\ \text{formamide}) - 600/\#\text{bp in duplex}$$

As an illustration of the above formula, using [N+] = [0.368] and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C. The T_m of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C.

Nucleic acids of the present invention may be maintained as DNA in any convenient cloning vector. In a preferred embodiment, clones are maintained in plasmid cloning/expression vector, such as pBluescript (Stratagene, La Jolla, CA), which is propagated in a suitable *E. coli* host cell.

P73 protein can be produced by using *in vitro* expression methods known in the art. For example, part or all of a DNA molecule, such as a DNA encoding the amino acid sequence SEQ ID NO:1 or SEQ ID NO:2, may be inserted into a plasmid vector adapted for expression in a bacterial cell, such as *E. coli*, or a eukaryotic cell, such as *Saccharomyces cerevisiae* or other yeast. In a preferred embodiment, a commercially available expression/secretion system can be used, whereby the recombinant protein is expressed and thereafter secreted from the host cell, to be easily purified from the surrounding medium. If expression/secretion vectors are not used, an alternative approach involves purifying the recombinant protein by affinity separation, such as by immunological interaction with antibodies that bind specifically to the recombinant protein or fusion proteins such as His tags. Such methods are commonly used by skilled practitioners.

The method of the invention for treating mammalian cells that over-express E6 comprises administering a therapeutically effective amount of p73 protein to the target cells. The administration of the p73 protein can be accomplished *via* several methods, including the exposing the target cell, i.e., the E6-over-expressing cell, to p73 protein, or exposing the target cell to a nucleic acid construct that expresses an appropriate p73 coding sequence.

Any method of administration of p73 (e.g, as a protein or as a nucleic acid encoding the protein) is appropriate as long as it results in increased levels of p73 protein within the target cell. The choice of method of administration will depend largely on the position of the target cells and the length of time the treatment is

needed. Target cells may be removed from the patient and treated *ex vivo*, and then reintroduced to the patient. Additionally, the treatment may be used in cell cultures for experimental purposes. In a preferred embodiment, the target cells comprise E6-over-expressing carcinomas. In a more preferred embodiment, the target cells are papilloma-virus positive cancers. In a most preferred embodiment, the target cells are HPV-positive carcinomas of the uterine cervix.

The administration of p73 protein to target cells can be accomplished by exposing the target cell to p73 protein. When the target cell are tumor cells within an animal, it is preferred that the protein is administered in a protected form to increase their stability cells One strategy of accomplishing this is to use liposomes. Liposomes are water-filled vesicles composed of several phospholipids layers surrounding an aqueous core with an outer shell capable of providing direction to specific target cells. Typically liposomes are composed of some combination of phosphatidylcholine, cholesterol, phosphatidylglycerol or other glycolipids or phospholipids (Hudson and Black, 1993, American Pharmacy NS33(5):23-24). Insoluble polymers composed of polyethylene may also be used to form a protective layer around the protein, inhibiting degradation while traveling to the target cell (Hudson and Black, 1993, American Pharmacy NS33(5):23-24). Another way to deliver p73 protein to target cells is to couple the protein to a target cell-specific monoclonal antibody. This approach allows the protein to be specifically delivered to the target cell and minimizes toxic effects on non-target cells (Houston, 1993, Current Opinion in Biotechnology 4:739-744).

In preferred embodiments, the p73 protein is administered to the target cell through the use of heterologous nucleic acids that will cause the protein to be synthesized within the target cell. These nucleic acids can be temporary residents in the target cell, such as expression plasmids, or they can be stably integrated into the genome of the target cell. Expression plasmids are particularly appropriate for experimental work with cell cultures, such as illustrated in Example 1. The construction of such plasmids and the transformation of target cells with them *in vitro* is well known to those of skill in the art of cell biology. Expression vectors suitable for p73 expression in mammalian cells are commercially available (Gene Therapy Systems, San Diego). Naked DNA and plasmids may be delivered to the target cells by several known means. The naked DNA may be transferred directly into the genetic material of the cells (Wolff et al., 1990, Science 247:1465-1468), the p73-encoding DNA may be delivered in liposomes (Ledley, 1987, J. Pediatrics 110:1) or proteoliposomes that contain viral envelope receptor proteins (Nicolau et al, 1983, Proc. Natl. Acad. Sci. U.S.A. 80:1068), or the p73-encoding DNA may be coupled to a polylysine-glycoprotein carrier complex.

For a longer lasting expression of p73 within target cells, viral vectors are preferred. A variety of viral vector may be used in this invention, included retroviral vectors such as the herpes simplex virus (U.S.Patent 5,288,641, incorporated herein by reference), Cytomegalovirus, murine leukemia virus (Blaese et al., 1995, Science 270:475-479) and similar as described by Miller (Miller, 1992, Curr. Top. Microbiol. Immunol. 158:1). Recombinant adeno-associated virus (AAV vectors)

such as those described by U.S. Patent No. 5,139,941 (which is incorporated herein by reference) and recombinant adenoviral vectors (He et al., 1998, PNAS 95:2509-2514, incorporated by reference herein) are particularly preferred. Also contemplated are recombinant lentivirus vectors such as a recombinant Human Immunodeficiency Virus (U.S. Patent No. 5,885,805; Blaese et al., 1995, Science 270:475-479; Onodera et al., 1998, J. of Virology 72:1769-1774) and Feline Immunodeficiency Virus. Often these vectors have been designed so that they are replication-defective, and the techniques to prepare such vectors are well known in the art (Ghosh-Choudhury and Graham, 1987, Biochem. Biophys. Res. Comm. 147:964-973; McGrory, W. J. et al., 1988, Virology 163:614-617; Gluzman et al., 1982 in Eukaryotic Viral Vectors (Gluzman, Y., Ed.) pp. 187-192, Cold Spring Harbor Press, Cold Spring Harbor, N.Y). It is also contemplated that viral vectors that are replication competent may be used to improve the efficacy of the treatment of solid tumors (Wildner et al., 1999, Gene Ther. 6:57-62).

The recombinant vector of the invention comprises a nucleic acid construct comprising a sequence encoding a p73 protein operably linked to an appropriate promoter and other expression-regulatory sequences. For treatment of cancer cells, a strong constitutive promoter, such as a cytomegalovirus promoter, a viral LTR, RSV or SV40 promoter, is preferred. In a preferred embodiment, a cytomegalovirus promoter is used. Additionally, promoters associated with genes that are expressed at high levels in mammalian cells, such as elongation factor-1 and actin, are also contemplated. It is particularly advantageous to use a viral-specific and

-regulated promoter to direct expression specifically in affected cancer cells. In a particularly preferred embodiment, the HPV-E6 promoter is used.

In a particularly preferred embodiment, a recombinant adenoviral vector is used to deliver the p73-expressing construct to the target cell. The use of adenoviral vectors for gene therapy is well known in the art (El-Deiry et al., 1993, Cell 75:817; Blogosklonny and El-Deiry, 1996, Int. J. Cancer 67:386-395; Prabhu et al., 1996, Clin Cancer Res. 2:1221-1230; Zeng et al., 1997, Int. J. Oncol. 11:221-226; Mitchell and El-Deiry, 1999, Cell Growth and Diff. 10:223-230; Meng et al., 1998, Clin. Cancer Res. 4:251-259; Blagosklonny and El-Deiry, 1998, Int. J. Cancer 75:933-940). In particular, an adenovirus vector has been used successfully to deliver p53 to target cells to treat lung cancer in human patients (Roth et al., 1996, Nature Med. 2:974 incorporated herein by reference; and U.S. Patent 5,747,469 incorporated herein by reference). It is contemplated that these protocols with simple variation that will be well known to those in the art can be used to administer the p73 protein to target cells in the invention. In a most preferred embodiment, therapeutically effective amounts of the viral vector are delivered to the cancers by direct injection.

The interchangeability of p53 and p73 in these methods arises from the high degree of similarity that these proteins have, both in structure and function. p53 and p73 have significant amino acid sequence similarities (Kaghad et al, 1997, Cell 90:809-818, incorporated by reference herein), particularly in the most conserved regions of p53: the transactivation, DNA binding and p53 oligomerization domains. A sequence similar to the MDM2-

binding domain is also present in p73. The residues in p53 often found mutated in tumors and shown to be required for DNA recognition are conserved and occupy identical positions in p73. The C-terminal domain of p73α shows homology to invertebrate p53 homologs. Finally the intron-exon organization of p73 is very similar to p53.

p53 and p73 are also functionally similar. Both display homotypic interactions, and p53 and p73β display significant mutual interactions (Kaghad et al, 1997, Cell 90:809-818). Both are inhibited by adenovirus E4ORF6 (Higashino et al., 1998, PNAS 95:15683-15687) and the MDM2 oncoprotein (Zeng et al., 1999, Mol. Cell. Biol. 19:327-3266; Dobbelstein et al., 1999, Oncogene 18:2101-2106). p73 function is inhibited by tumor-derived p53 mutants in mammalian cells in a manner similar to p53 (Di Como et al., 1999, Mol. Cell. Biol. 19:1438-1449). p73 regulates p53 target genes when p73 is over-expressed in cells (Zhu et al., 1998, Cancer Research 58:5061-5065; Jost et al., 1997, Nature 389:181-184). Finally, as a result of activation of p53-responsive genes, p73 can inhibit cell growth and induce apoptosis in a manner similar to p53.

Also provided with the invention are pharmaceutical compositions that can be used to treat mammalian cells with p73 *in vitro*, *in vivo* and *ex vivo*. The compositions comprise either p73 protein or nucleic acids encoding p73 protein. The pharmaceutical compositions of the invention are formulated in an appropriate "biologically acceptable medium". As used herein, "biologically acceptable medium" includes any and all solvents, dispersion media and the like which may be appropriate for the desired route of administration of

the pharmaceutical preparation, as exemplified in the preceding paragraph. The use of such media for pharmaceutically active substances is known in the art. Except insofar as any conventional media or agent is incompatible with the nucleic acid molecules or proteins to be administered, its use in the pharmaceutical preparation is contemplated.

The pharmaceutical preparation is formulated in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form, as used herein, refers to a physically discrete unit of the pharmaceutical preparation appropriate for the patient undergoing treatment. Each dosage should contain a quantity of active ingredient calculated to produce the desired effect in association with the selected pharmaceutical carrier. Procedures for determining the appropriate dosage unit are well known to those skilled in the art.

The pharmaceutical composition also can include various other components as additives or adjuncts. Exemplary pharmaceutically acceptable components or adjuncts which are employed in relevant circumstances include antioxidants, free radical scavenging agents, peptides, growth factors, antibiotics, bacteriostatic agents, immunosuppressives, anticoagulants, buffering agents, anti-inflammatory agents, anti-pyretics, time release binders, anaesthetics, steroids and corticosteroids. Such components can provide additional therapeutic benefit, act to effect the therapeutic action of the pharmaceutical composition, or act towards preventing any potential side effects which may be posed as a result of administration of the pharmaceutical composition. In certain circumstances, the p73 protein

or nucleic acid molecule can be employed as part of a pharmaceutical composition with other compounds (e.g., chemotherapeutic agents) intended to prevent or treat cancer or a related disorder.

The manner in which the pharmaceutical preparations are administered can vary. They can be administered by inhalation (e.g., in the form of an aerosol either nasally or using delivery articles of the type set forth in U.S. Patent No. 4,922,901 to Brooks et al.); topically (e.g., in lotion form or as a suppository); orally (e.g., in liquid form within a solvent such as an aqueous or non-aqueous liquid, or within a solid carrier); intravenously (e.g., within a dextrose or saline solution); as an infusion or injection (e.g., as a suspension or as an emulsion in a pharmaceutically acceptable liquid or mixture of liquids); intrathecally; intracerebro- ventricularly; or transdermally (e.g., using a transdermal patch). Exemplary methods for administering such compounds will be apparent to the skilled artisan. The administration of the pharmaceutical compositions of the present invention can be intermittent, or at a gradual, continuous, constant or controlled rate to a warm-blooded animal, (e.g., a mammal such as a mouse, rat, cat, rabbit, dog, pig, cow, or monkey); but advantageously is preferably administered to a human being. In addition, the time interval between administrations can vary. Administration preferably is such that the active ingredients of the pharmaceutical formulation contact the target cells, whether within or outside the body of a mammalian subject.

The appropriate dose of the compound is that amount effective to result in increased levels of p73

protein within the target cell. By "effective amount", "therapeutic amount" or "effective dose" is meant that amount sufficient to elicit the desired pharmacological or therapeutic effects, thus resulting in effective prevention or treatment of the disorder. Prevention of the disorder is manifested by delaying the onset of the symptoms of the disorder. Treatment of the disorder is manifested by a decrease in the symptoms associated with the disorder or an amelioration of the recurrence of the symptoms of the disorder.

The effective dose can vary, depending upon factors such as the condition of the patient, the severity of the symptoms of the disorder, and the manner in which the pharmaceutical composition is administered. The effective dose of compounds will of course differ from patient to patient but in general includes amounts starting where target cell growth is halted to where the target cell is killed. Dosages contemplated for use with the retroviral vector embodiment of the invention are those suggested in U.S. Patent 5,747,469 (incorporated herein by reference). One of ordinary skill in the art will know how to determine such doses without undue experimentation.

Kits with the components necessary to treat E6-over-expressing target cells by the method of the invention are also provided. In a preferred embodiment, the kit contains therapeutically effective amounts of the pharmaceutical preparation of the invention in a container. The pharmaceutical preparation in the kit may be comprised of p73 protein or a DNA construct encoding p73, preferably inserted into a vector for transforming cells. The p73 protein or p73 encoding viral vector may be in the form of a pharmaceutically acceptable sterile

solution such as sterile saline, dextrose solution or buffered solution. Alternatively, the p73 protein or p73 encoding viral vector can be lyophilized or desiccated. In this instance the kit may optionally further comprise a container of a pharmaceutically acceptable solution, (e.g., saline, dextrose solution, etc.), preferably sterile, to reconstitute the pharmaceutical preparation to form a solution for injection purposes. Optionally, instructions may be included in the kit. The kit may additionally comprise pharmaceutical preparations in containers for other therapies related to cancer treatment, such as chemotherapy.

The following example is provided to describe the invention in greater detail. It is intended to illustrate, not to limit, the invention.

EXAMPLE 1

p73α and p73β Suppress Growth and Induce Apoptosis in Human Papilloma Virus E6-Expressing Cancer Cells

Materials and Methods

Plasmids. The mammalian expression vector pCMV-neo-Bam (Baker et al., 1990, Science 249: 912-915) and the wild-type p53 expression vector SN3 (Baker et al., 1990, Science 249: 912-915) were obtained from Bert Vogelstein (Johns Hopkins University). Wild-type and mutant p73α and p73β plasmids (Jost et al.,1997, Nature 389: 191-194; incorporated herein by reference) were obtained from William G. Kaelin, Jr. (Dana Farber Cancer Institute). The HPV-E6 expression plasmid (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229; incorporated herein by reference) was obtained from Kathleen Cho (Johns Hopkins University).

Cell culture and transfection conditions. The mutant p53-expressing human colon adenocarcinoma cell line SW480 was maintained in culture as previously described (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229). The mutant p53-expressing human glioma cell lines U373 and H80 were obtained from Peter C. Phillips (The Children's Hospital of Philadelphia) and the wild-type p53-expressing human non-small cell lung cancer cell line H460 was obtained from Stephen B. Baylin (Johns Hopkins University). Mutant p53-expressing SKBr3 cells were obtained from American Type Culture Collection (Rockville, MD). SW480 cells were transfected using Lipofectin (BRL) as previously described (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229). At 20 hrs. following transfection, cells were harvested and protein lysates electrophoresed through 10% polyacrylamide gels and immunoblotted as previously described (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229). Analysis of p53 expression was performed using the anti-human p53 monoclonal antibody pAb1801 (Ab2; Oncogene Science). For detection of exogenous p73 protein expression, the anti-HA antibody was used as previously described (Jost et al.,1997, Nature 389: 191-194).

Adenovirus infections. The Ad-LacZ reagent (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229) was obtained from Bert Vogelstein. The HPV type 16 E6-expressing replication deficient adenovirus was prepared and titered as previously described (Satyamoorthy et al., 1997, Cancer Res. 57: 1873-1876; Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229). Briefly, the CMV promoter-driven HPV type 16 E6 cDNA was inserted into an E3-deleted adenovirus by homologous recombination to generate E1 and E3 deleted replication

defective Ad-E6 adenovirus (Satyamoorthy et al., 1997, Cancer Res. 57: 1873-1876; incorporated herein by reference). The cloned HPV-E6 DNA sequence was verified and expression of HPV-E6 was verified by Northern blotting of total RNA derived from Ad-E6 versus Ad-LacZ infected cells. Cells were infected using an MOI of 50 as previously described (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229). Infection of SW480 cells using Ad-LacZ at an MOI of 50 followed by X-gal staining revealed greater than 99% infectivity.

Colony suppression assays. Transfections were carried out as described above except that the tumor suppressive (p53 or p73) or control (pCMV-neo-Bam) plasmid represented 80% of the total DNA and the degrading (pCMV-E6) or control (pCMV-neo-Bam) represented the remaining 20% of the total DNA. At 24 hrs following transfection, G418 selection was begun using 1 mg/ml as a final concentration. Selection was continued for 7-12 days and colony growth was analyzed as previously described (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229).

TUNEL assays. At 48 hrs following transfection, cells were formalin fixed and the extent of apoptosis was assessed by nicked-end labeling using the Apotag kit (Oncor) followed by analysis using fluorescence microscopy.

Results

HPV-E6 targets both wild-type and mutant p53 protein for degradation. Using an E6-expressing adenovirus (Ad-E6) a panel of human cancer cells derived from different tissues and containing either endogenous wild-type or mutant p53 were infected (Fig. 1). As

compared with Ad-LacZ infected cells, E6-expressing cells expressed substantially reduced levels of either wild-type or mutant p53 protein (Fig. 1, compare even to odd lanes). Thus the HPV-E6 protein targets both wild-type and mutant p53 for degradation. E6 does not target the cell cycle regulatory proteins pRb, p21, cyclin E or p27 for degradation (Fig. 1 lower panels). The phosphorylation state of Rb in Ad-E6 was not altered as compared to Ad-LacZ infected cells (Fig. 1 lower panels), regardless of the p53 status of the cells.

p73 is resistant to HPV E6-dependent proteolysis. Because p53 is degraded in HPV-E6 expressing cancer cells, such cells are not ideally suited for gene replacement therapy (Prabhu et al., 1996, Clin. Cancer Res. 2: 1221-1229). In HPV E6-expressing cancer cells, p53 is degraded while exogenous p73 is resistant to E6-targeting to the proteasome (Fig. 2). The resistance of p73 to E6-dependent proteolysis was observed with p73α, p73αm, p73β, or p73βm. This observation suggested that p73 is a candidate for gene replacement in E6-expressing cancer cells.

p73β induces apoptosis and suppresses growth in HPV E6-expressing human cancer cells. p73β was previously found to be a potent activator of p53-dependent gene expression (Kaghad et al., 1997, Cell 90: 809-819; Jost et al.,1997, Nature 389: 191-194). p73β in colony suppression assays in the absence or presence of E6-expression. Whereas p53 failed to inhibit the growth of E6-expressing cancer cells, p73β was found to be a potent growth suppressor. Transfection studies revealed

that p73α was a less potent suppressor of growth of SW480 cancer cells either in the absence or presence of HPV-E6.

This could not be explained by dominant negative inhibition of p73α by the endogenous p53 mutant in SW480 cells because it was previously shown that p73α shows negligible interaction with p53 (Kaghad et al., 1997, Cell 90: 809-819).

p73β has been previously shown to be an inducer of apoptosis, similar to p53 (Jost et al.,1997, Nature 389: 191-194). Whereas p53-dependent apoptosis was inhibited in E6-expressing cells, p73β was still capable of inducing apoptosis similar to what is observed in the absence of E6. Therefore the colony suppression phenotype observed following p53 or p73β expression in the presence or absence of E6 can be explained by their ability to induce apoptosis under these conditions. These results suggest that the p73β-dependent suppression of growth of HPV E6-expressing cancer cells occurs through an apoptotic mechanism.

The present invention is not limited to the embodiments described and exemplified above, but is capable of variation and modification without departure from the scope of the appended claims.

What is claimed:

1. A method of inducing apoptosis in an E6-expressing cell, comprising administering to the cell an amount of p73 protein effective to induce the apoptosis.

2. The method of claim 1 wherein the p53 protein is administered as a DNA construct comprising an expressible sequence that encodes the p73 protein.

3. The method of claim 2, wherein the DNA construct is operably inserted into a viral vector.

4. The method of claim 3, wherein the viral vector is selected from the group consisting of adenoviral vectors, HIV vectors, FIV vectors, herpes viral vectors, adeno-associated vectors and cytovegaviral vectors.

5. The method of claim 1, wherein the cell is a cancerous cell.

6. The method of claim 1, wherein the cell is infected with Human papilloma virus.

7. The method of claim 1, wherein the cell is a cultured cell.

8. The method of claim 1, wherein the cell is obtained from the body of a living organism, the administering is performed *ex vivo*, and the cell is returned to the living organism.

9. The method of claim 1, wherein the cell is disposed within a living organism and the administering is performed *in vivo*.

10. The method of claim 1, wherein the cell is obtained from a species selected from the group consisting of *Homo sapiens*, *Mus musculus* and *Chlorocebus aethiops*.

11. The method of claim 1, wherein the p73 protein is p73α or p73β.

12. The method of claim 11, wherein the protein comprises a sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:2.

13. The method of claim 2, wherein the DNA construct comprises more than 50 nucleotides of SEQ ID NO:3.

14. A pharmaceutical preparation for treatment of cancers associated with E6 over-expression, comprising p73 protein associated with a delivery vehicle for delivering the preparation to cancer cells.

15. The pharmaceutical preparation of claim 14, wherein the p73 protein is p73α or p73β.

16. The pharmaceutical preparation of claim 15, wherein the protein comprises a sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:2.

17. The pharmaceutical preparation of claim 14, which further comprises at least one additional active ingredient for treatment of cancer.

18. A kit comprising a container containing one or more dosage units of the pharmaceutical composition of claim 14.

19. The kit of claim 18, which further comprises at least one additional pharmaceutical agent for treatment of cancer.

20. A pharmaceutical preparation for treatment of cancers associated with E6 over-expression, comprising an expressible DNA construct encoding p73, associated with a delivery vehicle for delivering the preparation to cancer cells.

21. The pharmaceutical preparation of claim 20, wherein the DNA construct is operably inserted into a vector for transforming cells.

22. The pharmaceutical preparation of claim 21, wherein the vector is a viral vector selected from the group consisting of adenoviral vectors, HIV vectors, FIV vectors, herpes viral vectors, adeno-associated vectors and cytovegaviral vectors.

23. A kit comprising a container containing one or more dosage units of the pharmaceutical composition of claim 20.

24. The kit of claim 23, which further comprises at least one additional pharmaceutical agent for treatment of cancer.

25. An apoptotic, E6-expressing transgenic cell comprising a heterologous, expressible DNA construct encoding p73.

26. The cell of claim 25, obtained from a cultured cell line.

27. The cell of claim 25, disposed within a living organism.

28. The cell of claim 25, from a species selected from the group consisting of *Homo sapiens*, *Mus musculus* and *Chlorocebus aethiops*.

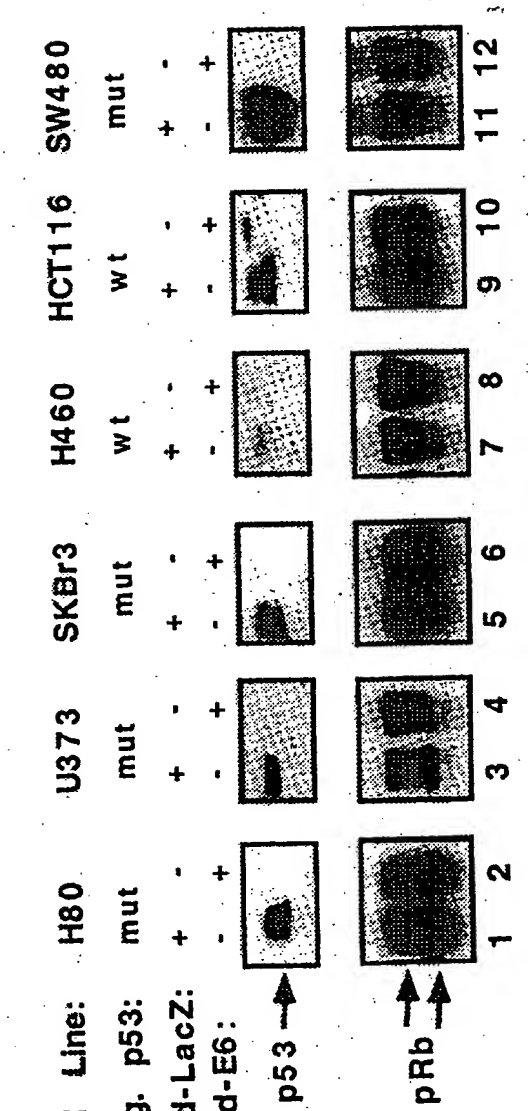
Cell Line: H80 U373 SKBr3 H460 HCT116 SW480
Endog. p53: mut mut mut wt wt mut
Ad-LacZ: + - + - + - + - + - + -
Ad-E6: - + - + - + - + - + - +
p53
pRb
1 2 3 4 5 6 7 8 9 10 11 12

Figure 1

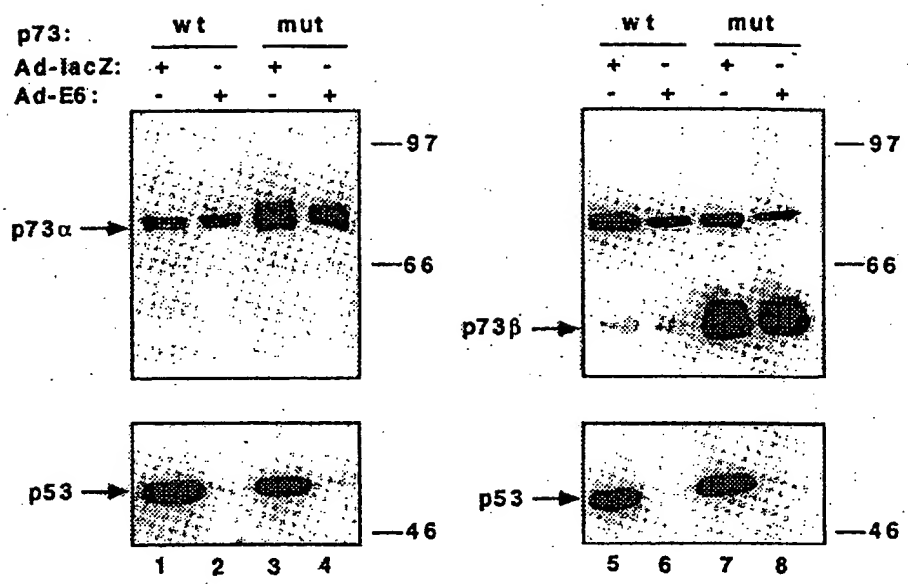

Figure 2

SEQUENCE LISTING

<110> El-Deiry, Wafik

<120> Compositions and Methods for Inducing Apoptosis in E6-Expressing Cells

<130> Penn K-1824

<150> 60/090,526
<151> 1998-06-24

<160> 3

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 636
<212> PRT
<213> Homo sapiens

<300>
<301> Kaghad, M., Bonnet, H., Yang, A., Creancier, L., B
<302> Monoallelecally expressed gene related to p53 at 1
<303> Cell
<304> 90
<305> 4
<306> 809-819
<307> 1997-08-22
<308> Genbank Y11416
<309> 1997-09-02

<400> 1

```
Met Ala Gln Ser Thr Ala Thr Ser Pro Asp Gly Gly Thr Thr Phe Glu
 1               5                  10                  15
His Leu Trp Ser Ser Leu Glu Pro Asp Ser Thr Tyr Phe Asp Leu Pro
            20                  25                  30
Gln Ser Ser Arg Gly Asn Asn Glu Val Val Gly Gly Thr Asp Ser Ser
        35                  40                  45
Met Asp Val Phe His Leu Glu Gly Met Thr Thr Ser Val Met Ala Gln
    50                  55                  60
Phe Asn Leu Leu Ser Ser Thr Met Asp Gln Met Ser Ser Arg Ala Ala
65                  70                  75                  80
Ser Ala Ser Pro Tyr Thr Pro Glu His Ala Ala Ser Val Pro Thr His
                85                  90                  95
Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Thr Met Ser Pro Ala
            100                 105                 110
Pro Val Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro His His Phe Glu
        115                 120                 125
Val Thr Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr
    130                 135                 140
Ser Pro Leu Leu Lys Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro
145                 150                 155                 160
Ile Gln Ile Lys Val Ser Thr Pro Pro Pro Pro Gly Thr Ala Ile Arg
                165                 170                 175
Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr Asp Val Val Lys
            180                 185                 190
Arg Cys Pro Asn His Glu Leu Gly Arg Asp Phe Asn Glu Gly Gln Ser
```

```
                195                 200                 205
Ala Pro Ala Ser His Leu Ile Arg Val Glu Gly Asn Asn Leu Ser Gln
        210                 215                 220
Tyr Val Asp Asp Pro Val Thr Gly Arg Gln Ser Val Val Val Pro Tyr
225                 230                 235                 240
Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Ile Leu Tyr Asn Phe
                245                 250                 255
Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu
            260                 265                 270
Ile Ile Ile Thr Leu Glu Met Arg Asp Gly Gln Val Leu Gly Arg Arg
        275                 280                 285
Ser Phe Glu Gly Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala
    290                 295                 300
Asp Glu Asp His Tyr Arg Glu Gln Gln Ala Leu Asn Glu Ser Ser Ala
305                 310                 315                 320
Lys Asn Gly Ala Ala Ser Lys Arg Ala Phe Lys Gln Ser Pro Pro Ala
                325                 330                 335
Val Pro Ala Leu Gly Ala Gly Val Lys Lys Arg Arg His Gly Asp Glu
            340                 345                 350
Asp Thr Tyr Tyr Leu Gln Val Arg Gly Arg Glu Asn Phe Glu Ile Leu
        355                 360                 365
Met Lys Leu Lys Glu Ser Leu Glu Leu Met Glu Leu Val Pro Gln Pro
    370                 375                 380
Leu Val Asp Ser Tyr Arg Gln Gln Gln Gln Leu Leu Gln Arg Pro Ser
385                 390                 395                 400
His Leu Gln Pro Pro Ser Tyr Gly Pro Val Leu Ser Pro Met Asn Lys
                405                 410                 415
Val His Gly Gly Met Asn Lys Leu Pro Ser Val Asn Gln Leu Val Gly
            420                 425                 430
Gln Pro Pro Pro His Ser Ser Ala Ala Thr Pro Asn Leu Gly Pro Val
        435                 440                 445
Gly Pro Gly Met Leu Asn Asn His Gly His Ala Val Pro Ala Asn Gly
    450                 455                 460
Glu Met Ser Ser Ser His Ser Ala Gln Ser Met Val Ser Gly Ser His
465                 470                 475                 480
Cys Thr Pro Pro Pro Pro Tyr His Ala Asp Pro Ser Leu Val Ser Phe
                485                 490                 495
Leu Thr Gly Leu Gly Cys Pro Asn Cys Ile Glu Tyr Phe Thr Ser Gln
            500                 505                 510
Gly Leu Gln Ser Ile Tyr His Leu Gln Asn Leu Thr Ile Glu Asp Leu
        515                 520                 525
Gly Ala Leu Lys Ile Pro Glu Gln Tyr Arg Met Thr Ile Trp Arg Gly
    530                 535                 540
Leu Gln Asp Leu Lys Gln Gly His Asp Tyr Ser Thr Ala Gln Gln Leu
545                 550                 555                 560
Leu Arg Ser Ser Asn Ala Ala Thr Ile Ser Ile Gly Gly Ser Gly Glu
                565                 570                 575
Leu Gln Arg Gln Arg Val Met Glu Ala Val His Phe Arg Val Arg His
            580                 585                 590
Thr Ile Thr Ile Pro Asn Arg Gly Gly Pro Gly Gly Gly Pro Asp Glu
        595                 600                 605
Trp Ala Asp Phe Gly Phe Asp Leu Pro Asp Cys Lys Ala Arg Lys Gln
    610                 615                 620
Pro Ile Lys Glu Glu Phe Thr Glu Ala Glu Ile His
625                 630                 635
```

<210> 2
<211> 499
<212> PRT
<213> Homo sapiens

<300>
<301> Kaghad et al.,
<302> Monoallelically expressed gene related to p53 at 1
<303> Cell
<304> 90
<305> 4
<306> 809-819
<307> 1997-08-22
<308> Genbank Y11416
<309> 1997-09-02

<400> 2

```
Met Ala Gln Ser Thr Ala Thr Ser Pro Asp Gly Gly Thr Thr Phe Glu
 1               5                  10                  15
His Leu Trp Ser Ser Leu Glu Pro Asp Ser Thr Tyr Phe Asp Leu Pro
            20                  25                  30
Gln Ser Ser Arg Gly Asn Asn Glu Val Val Gly Gly Thr Asp Ser Ser
        35                  40                  45
Met Asp Val Phe His Leu Glu Gly Met Thr Thr Ser Val Met Ala Gln
    50                  55                  60
Phe Asn Leu Leu Ser Ser Thr Met Asp Gln Met Ser Ser Arg Ala Ala
65                  70                  75                  80
Ser Ala Ser Pro Tyr Thr Pro Glu His Ala Ala Ser Val Pro Thr His
                85                  90                  95
Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Thr Met Ser Pro Ala
            100                 105                 110
Pro Val Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro His His Phe Glu
        115                 120                 125
Val Thr Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr
    130                 135                 140
Ser Pro Leu Leu Lys Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro
145                 150                 155                 160
Ile Gln Ile Lys Val Ser Thr Pro Pro Pro Pro Gly Thr Ala Ile Arg
                165                 170                 175
Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr Asp Val Val Lys
            180                 185                 190
Arg Cys Pro Asn His Glu Leu Gly Arg Asp Phe Asn Glu Gly Gln Ser
        195                 200                 205
Ala Pro Ala Ser His Leu Ile Arg Val Glu Gly Asn Asn Leu Ser Gln
    210                 215                 220
Tyr Val Asp Asp Pro Val Thr Gly Arg Gln Ser Val Val Val Pro Tyr
225                 230                 235                 240
Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Ile Leu Tyr Asn Phe
                245                 250                 255
Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu
            260                 265                 270
Ile Ile Ile Thr Leu Glu Met Arg Asp Gly Gln Val Leu Gly Arg Arg
        275                 280                 285
Ser Phe Glu Gly Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala
    290                 295                 300
Asp Glu Asp His Tyr Arg Glu Gln Gln Ala Leu Asn Glu Ser Ser Ala
305                 310                 315                 320
Lys Asn Gly Ala Ala Ser Lys Arg Ala Phe Lys Gln Ser Pro Pro Ala
                325                 330                 335
Val Pro Ala Leu Gly Ala Gly Val Lys Lys Arg Arg His Gly Asp Glu
            340                 345                 350
Asp Thr Tyr Tyr Leu Gln Val Arg Gly Arg Glu Asn Phe Glu Ile Leu
        355                 360                 365
Met Lys Leu Lys Glu Ser Leu Glu Leu Met Glu Leu Val Pro Gln Pro
    370                 375                 380
Leu Val Asp Ser Tyr Arg Gln Gln Gln Gln Leu Leu Gln Arg Pro Ser
```

385 390 395 400
His Leu Gln Pro Pro Ser Tyr Gly Pro Val Leu Ser Pro Met Asn Lys
405 410 415
Val His Gly Gly Met Asn Lys Leu Pro Ser Val Asn Gln Leu Val Gly
420 425 430
Gln Pro Pro Pro His Ser Ser Ala Ala Thr Pro Asn Leu Gly Pro Val
435 440 445
Gly Pro Gly Met Leu Asn Asn His Gly His Ala Val Pro Ala Asn Gly
450 455 460
Glu Met Ser Ser Ser His Ser Ala Gln Ser Met Val Ser Gly Ser His
465 470 475 480
Cys Thr Pro Pro Pro Pro Tyr His Ala Asp Pro Ser Leu Val Arg Thr
485 490 495
Trp Gly Pro

<210> 3
<211> 1594
<212> DNA
<213> Homo sapiens

<300>
<301> Kaghad et al.
<302> Monoallelically expressed gene related to p53 at 1
<303> Cell
<304> 90
<305> 4
<306> 809-819
<307> 1997-08-22
<308> Genbank Y11416
<309> 1997-09-02

<400> 3
atggcccagt ccaccgccac ctcccctgat gggggcacca cgtttgagca cctctggagc 60
tctctggaac cagacagcac ctacttcgac cttccccagt caagccgggg gaataatgag 120
gtggtgggcg gaacggattc cagcatggac gtcttccacc tggagggcat gactacatct 180
gtcatggccc agttcaatct gctgagcagc accatggacc agatgagcag ccgcgcggcc 240
tcggccagcc cctacacccc agagcacgcc gccagcgtgc ccacccactc gccctacgca 300
caacccagct ccaccttcga caccatgtcg ccggcgcctg tcatcccctc caacaccgac 360
taccccggac cccaccactt tgaggtcact ttccagcagt ccagcacggc caagtcagcc 420
acctggacgt actccccgct cttgaagaaa ctctactgcc agatcgccaa gacatgcccc 480
atccagatca aggtgtccac cccgccaccc ccaggcactg ccatccgggc catgcctgtt 540
tacaagaaag cggagcacgt gaccgacgtc gtgaaacgct gccccaacca cgagctcggg 600
agggacttca acgaaggaca gtctgctcca gccagccacc tcatccgcgt ggaaggcaat 660
aatctctcgc agtatgtgga tgaccctgtc accggcaggc agagcgtcgt ggtgccctat 720
gagccaccac aggtggggac ggaattcacc accatcctgt acaacttcat gtgtaacagc 780
agctgtgtag ggggcatgaa ccggcggccc atcctcatca tcatcaccct ggagatgcgg 840
gatgggcagg tgctgggccg ccggtccttt gagggccgca tctgcgcctg tcctggccgc 900
gaccgaaaag ctgatgagga ccactaccgg gagcagcagg ccctgaacga gagctccgcc 960
aagaacgggg ccgccagcaa gcgtgccttc aagcagagcc cccctgccgt ccccgccctt 1020
ggtgccggtg tgaagaagcg gcggcatgga gacgaggaca cgtactacct tcaggtgcga 1080
ggccgggaga actttgagat cctgatgaag ctgaaagaga gcctggagct gatggagttg 1140
gtgccgcagc cactggtgga ctcctatcgg cagcagcagc agctcctaca gaggccgagt 1200
cacctacagc ccccgtccta cgggccggtc ctctcgccca tgaacaaggt gcacgggggc 1260
atgaacaagc tgccctccgt caaccagctg gtgggccagc ctcccccgca cagttcggca 1320
gctacaccca acctggggcc cgtgggcccc gggatgctca acaaccatgg ccacgcagtg 1380
ccagccaacg gcgagatgag cagcagccac agcgcccagt ccatggtctc ggggtcccac 1440
tgcactccgc caccccccta ccacgccgac cccagcctcg tcagtttttt aacaggattg 1500
gggtgtccaa actgcatcga gtatttcacc tcccaagggt tacagagcat ttaccacctg 1560
cagaacctga ccattgagga cctgggggcc ctga 1594

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/14057

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 38/00, 48/00; C12N 15/85
US CL : 514/2, 44; 424/93.21, 93.3; 435/455, 456

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 514/2, 44; 424/93.21, 93.3; 435/455, 456

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS: MEDLINE, BIOSIS, EMBASE, CAPLUS, BIOTECHDS
search terms: p73, e6, papilloma, aethiops
SEQUENCE SEARCH; SEQ ID NOS 1-3

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, Y	Database Medline, AN 1998290825, PRABHU et al. p73beta, unlike p53, suppresses growth and induces apoptosis of human papillomavirus E6-expressing cancer cells. International Journal of Oncology. July 1998. Vol. 13. pages 5-9, abstract only.	1-28
A	JOST et al. p73 is a human p53-related protein that can induce apoptosis. Nature. 11 September 1997. Vol. 389. pages 191-194, entire document.	1-28
A	KAGHAD et al. Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. Cell. August 1997. Vol 90. pages 809-819, entire document.	1-28

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 07 SEPTEMBER 1999	Date of mailing of the international search report 28 OCT 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer RICHARD SCHNIZER Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/14057

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, A	MARIN et al. Viral oncoproteins discriminate between p53 and the p53 homologue p73. Molecular and Cellular Biology. November 1998. Vol. 16, pages 6316-6324, entire document.	1-28
A	OREN. Lonely no more: p53 finds its kin in a tumor suppressor haven. Cell. September 1997. Vol 90. Pages 829-832, entire document.	1-28

SEQ ID No. 6 of Serial No. 09/125,005 (SR-p70)

<400> 6

```
Met Ala Gln Ser Thr Ala Thr Ser Pro Asp Gly Gly Thr Thr Phe Glu
 1               5                  10                  15
His Leu Trp Ser Ser Leu Glu Pro Asp Ser Thr Tyr Phe Asp Leu Pro
            20                  25                  30
Gln Ser Ser Arg Gly Asn Asn Glu Val Val Gly Gly Thr Asp Ser Ser
        35                  40                  45
Met Asp Val Phe His Leu Glu Gly Met Thr Thr Ser Val Met Ala Gln
    50                  55                  60
Phe Asn Leu Leu Ser Ser Thr Met Asp Gln Met Ser Ser Arg Ala Ala
65                  70                  75                  80
Ser Ala Ser Pro Tyr Thr Pro Glu His Ala Ala Ser Val Pro Thr His
                85                  90                  95
Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Thr Met Ser Pro Ala
            100                 105                 110
Pro Val Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro His His Phe Glu
        115                 120                 125
Val Thr Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr
    130                 135                 140
Ser Pro Leu Leu Lys Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro
145                 150                 155                 160
Ile Gln Ile Lys Val Ser Thr Pro Pro Pro Pro Gly Thr Ala Ile Arg
                165                 170                 175
Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr Asp Val Val Lys
            180                 185                 190
Arg Cys Pro Asn His Glu Leu Gly Arg Asp Phe Asn Glu Gly Gln Ser
        195                 200                 205
Ala Pro Ala Ser His Leu Ile Arg Val Glu Gly Asn Asn Leu Ser Gln
    210                 215                 220
Tyr Val Asp Asp Pro Val Thr Gly Arg Gln Ser Val Val Val Pro Tyr
225                 230                 235                 240
Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Ile Leu Tyr Asn Phe
                245                 250                 255
Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu
            260                 265                 270
Ile Ile Ile Thr Leu Glu Met Arg Asp Gly Gln Val Leu Gly Arg Arg
        275                 280                 285
Ser Phe Glu Gly Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala
    290                 295                 300
Asp Glu Asp His Tyr Arg Glu Gln Gln Ala Leu Asn Glu Ser Ser Ala
305                 310                 315                 320
Lys Asn Gly Ala Ala Ser Lys Arg Ala Phe Lys Gln Ser Pro Pro Ala
                325                 330                 335
Val Pro Ala Leu Gly Ala Gly Val Lys Lys Arg Arg His Gly Asp Glu
            340                 345                 350
Asp Thr Tyr Tyr Leu Gln Val Arg Gly Arg Glu Asn Phe Glu Ile Leu
        355                 360                 365
Met Lys Leu Lys Glu Ser Leu Glu Leu Met Glu Leu Val Pro Gln Pro
    370                 375                 380
Leu Val Asp Ser Tyr Arg Gln Gln Gln Gln Leu Leu Gln Arg Pro Ser
385                 390                 395                 400
His Leu Gln Pro Pro Ser Tyr Gly Pro Val Leu Ser Pro Met Asn Lys
                405                 410                 415
Val His Gly Gly Met Asn Lys Leu Pro Ser Val Asn Gln Leu Val Gly
            420                 425                 430
Gln Pro Pro Pro His Ser Ser Ala Ala Thr Pro Asn Leu Gly Pro Val
        435                 440                 445
Gly Pro Gly Met Leu Asn Asn His Gly His Ala Val Pro Ala Asn Gly
    450                 455                 460
Glu Met Ser Ser Ser His Ser Ala Gln Ser Met Val Ser Gly Ser His
465                 470                 475                 480
Cys Thr Pro Pro Pro Pro Tyr His Ala Asp Pro Ser Leu Val Ser Phe
                485                 490                 495
Leu Thr Gly Leu Gly Cys Pro Asn Cys Ile Glu Tyr Phe Thr Ser Gln
            500                 505                 510
Gly Leu Gln Ser Ile Tyr His Leu Gln Asn Leu Thr Ile Glu Asp Leu
        515                 520                 525
Gly Ala Leu Lys Ile Pro Glu Gln Tyr Arg Met Thr Ile Trp Arg Gly
    530                 535                 540
Leu Gln Asp Leu Lys Gln Gly His Asp Tyr Ser Thr Ala Gln Gln Leu
545                 550                 555                 560
Leu Arg Ser Ser Asn Ala Ala Thr Ile Ser Ile Gly Gly Ser Gly Glu
                565                 570                 575
Leu Gln Arg Gln Arg Val Met Glu Ala Val His Phe Arg Val Arg His
            580                 585                 590
Thr Ile Thr Ile Pro Asn Arg Gly Gly Pro Gly Gly Gly Pro Asp Glu
        595                 600                 605
Trp Ala Asp Phe Gly Phe Asp Leu Pro Asp Cys Lys Ala Arg Lys Gln
    610                 615                 620
Pro Ile Lys Glu Glu Phe Thr Glu Ala Glu Ile His
625                 630                 635
```

<400> 1

Met Ala Gln Ser Thr Ala Thr Ser Pro Asp Gly Gly Thr Thr Phe Glu
1 5 10 15
His Leu Trp Ser Ser Leu Glu Pro Asp Ser Thr Tyr Phe Asp Leu Pro
20 25 30
Gln Ser Ser Arg Gly Asn Asn Glu Val Val Gly Gly Thr Asp Ser Ser
35 40 45
Met Asp Val Phe His Leu Glu Gly Met Thr Thr Ser Val Met Ala Gln
50 55 60
Phe Asn Leu Leu Ser Ser Thr Met Asp Gln Met Ser Ser Arg Ala Ala
65 70 75 80
Ser Ala Ser Pro Tyr Thr Pro Glu His Ala Ala Ser Val Pro Thr His
85 90 95
Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Thr Met Ser Pro Ala
100 105 110
Pro Val Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro His His Phe Glu
115 120 125
Val Thr Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr
130 135 140
Ser Pro Leu Leu Lys Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro
145 150 155 160
Ile Gln Ile Lys Val Ser Thr Pro Pro Pro Pro Gly Thr Ala Ile Arg
165 170 175
Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr Asp Val Val Lys
180 185 190
Arg Cys Pro Asn His Glu Leu Gly Arg Asp Phe Asn Glu Gly Gln Ser
195 200 205
Ala Pro Ala Ser His Leu Ile Arg Val Glu Gly Asn Asn Leu Ser Gln
210 215 220
Tyr Val Asp Asp Pro Val Thr Gly Arg Gln Ser Val Val Val Pro Tyr
225 230 235 240
Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Ile Leu Tyr Asn Phe
245 250 255
Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu
260 265 270
Ile Ile Ile Thr Leu Glu Met Arg Asp Gly Gln Val Leu Gly Arg Arg
275 280 285
Ser Phe Glu Gly Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala
290 295 300
Asp Glu Asp His Tyr Arg Glu Gln Gln Ala Leu Asn Glu Ser Ser Ala
305 310 315 320
Lys Asn Gly Ala Ala Ser Lys Arg Ala Phe Lys Gln Ser Pro Pro Ala
325 330 335
Val Pro Ala Leu Gly Ala Gly Val Lys Lys Arg Arg His Gly Asp Glu
340 345 350
Asp Thr Tyr Tyr Leu Gln Val Arg Gly Arg Glu Asn Phe Glu Ile Leu
355 360 365
Met Lys Leu Lys Glu Ser Leu Glu Leu Met Glu Leu Val Pro Gln Pro
370 375 380
Leu Val Asp Ser Tyr Arg Gln Gln Gln Gln Leu Leu Gln Arg Pro Ser
385 390 395 400
His Leu Gln Pro Pro Ser Tyr Gly Pro Val Leu Ser Pro Met Asn Lys
405 410 415
Val His Gly Gly Met Asn Lys Leu Pro Ser Val Asn Gln Leu Val Gly
420 425 430
Gln Pro Pro Pro His Ser Ser Ala Ala Thr Pro Asn Leu Gly Pro Val
435 440 445
Gly Pro Gly Met Leu Asn Asn His Gly His Ala Val Pro Ala Asn Gly
450 455 460
Glu Met Ser Ser Ser His Ser Ala Gln Ser Met Val Ser Gly Ser His
465 470 475 480
Cys Thr Pro Pro Pro Pro Tyr His Ala Asp Pro Ser Leu Val Ser Phe
485 490 495
Leu Thr Gly Leu Gly Cys Pro Asn Cys Ile Glu Tyr Phe Thr Ser Gln
500 505 510
Gly Leu Gln Ser Ile Tyr His Leu Gln Asn Leu Thr Ile Glu Asp Leu
515 520 525
Gly Ala Leu Lys Ile Pro Glu Gln Tyr Arg Met Thr Ile Trp Arg Gly
530 535 540
Leu Gln Asp Leu Lys Gln Gly His Asp Tyr Ser Thr Ala Gln Gln Leu
545 550 555 560
Leu Arg Ser Ser Asn Ala Ala Thr Ile Ser Ile Gly Gly Ser Gly Glu
565 570 575
Leu Gln Arg Gln Arg Val Met Glu Ala Val His Phe Arg Val Arg His
580 585 590
Thr Ile Thr Ile Pro Asn Arg Gly Gly Pro Gly Gly Gly Pro Asp Glu
595 600 605
Trp Ala Asp Phe Gly Phe Asp Leu Pro Asp Cys Lys Ala Arg Lys Gln
610 615 620
Pro Ile Lys Glu Glu Phe Thr Glu Ala Glu Ile His
625 630 635

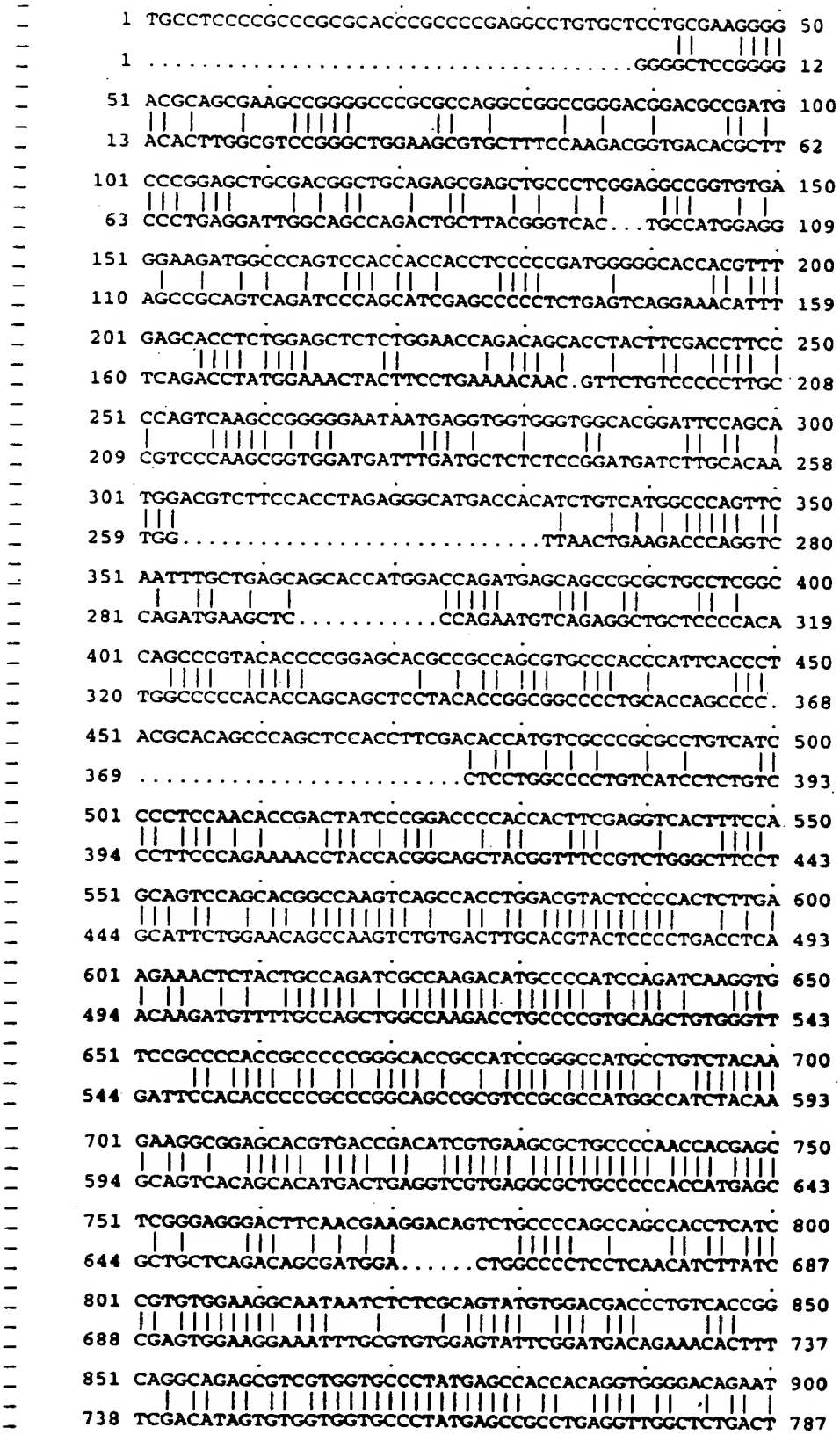

FIG. 1A

FIG.1

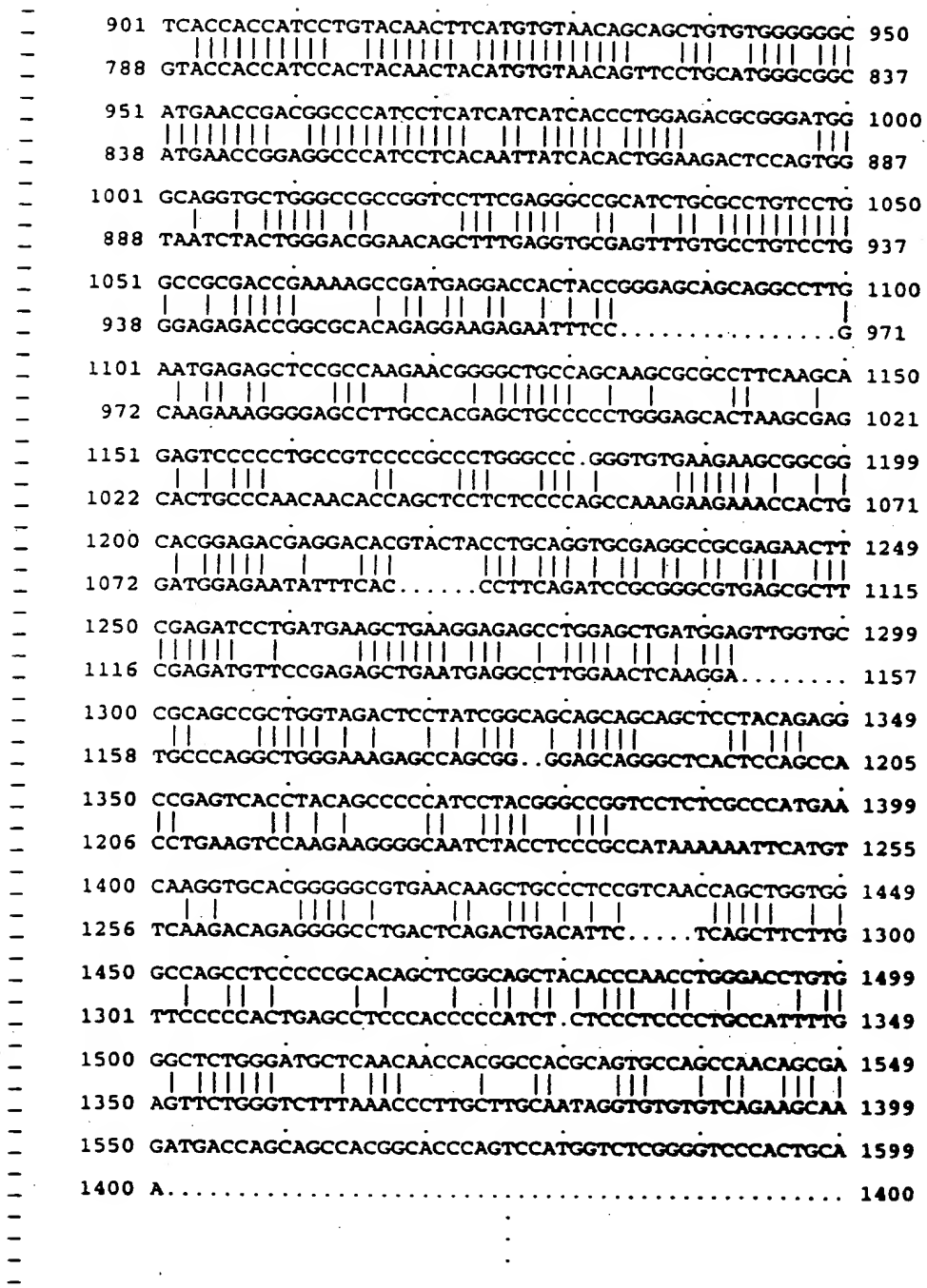

FIG. 1B

~~FIG.1 cont.~~

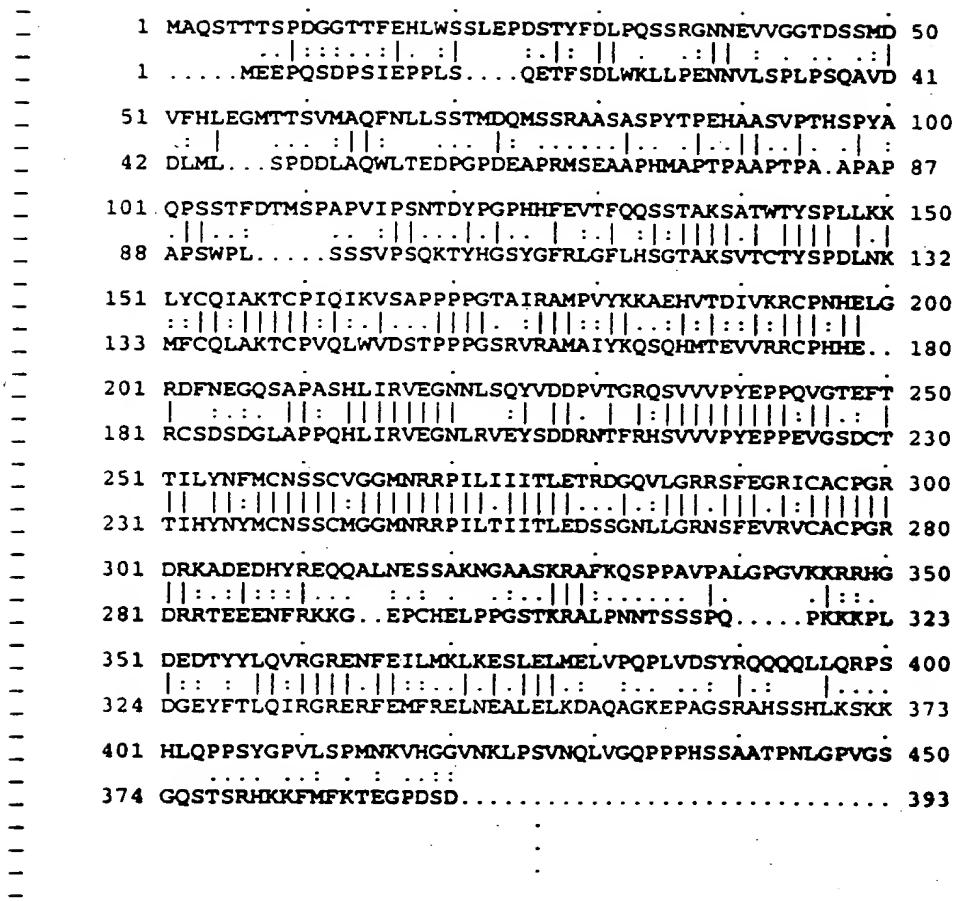
1 MAQSTTTSPDGGTTFEHLWSSLEPDSTYFDLPQSSRGNNEVVGGTDSSMD 50
1 .....MEEPQSDPSIEPPLS....QETFSDLWKLLPENNVLSPLPSQAVD 41
51 VFHLEGMTTSVMAQFNLLSSTMDQMSSRAASASPYTPEHAASVPTHSPYA 100
42 DLML...SPDDLAQWLTEDPGPDEAPRMSEAAPHMAPTPAAPTPA.APAP 87
101 QPSSTFDTMSPAPVIPSNTDYPGPHHFEVTFQQSSTAKSATWTYSPLLKK 150
88 APSWPL.....SSSVPSQKTYHGSYGFRLGFLHSGTAKSVTCTYSPDLNK 132
151 LYCQIAKTCPIQIKVSAPPPPGTAIRAMPVYKKAEHVTDIVKRCPNHELG 200
133 MFCQLAKTCPVQLWVDSTPPPGSRVRAMAIYKQSQHMTEVVRRCPHHE.. 180
201 RDFNEGQSAPASHLIRVEGNNLSQYVDDPVTGRQSVVVPYEPPQVGTEFT 250
181 RCSDSDGLAPPQHLIRVEGNLRVEYSDDRNTFRHSVVVPYEPPEVGSDCT 230
251 TILYNFMCNSSCVGGMNRRPILIIITLETRDGQVLGRRSFEGRICACPGR 300
231 TIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGR 280
301 DRKADEDHYREQQALNESSAKNGAASKRAFKQSPPAVPALGPGVKKRRHG 350
281 DRRTEEENFRKKG..EPCHELPPGSTKRALPNNTSSSPQ.....PKKKPL 323
351 DEDTYYLQVRGRENFEILMKLKESLELMELVPQPLVDSYRQQQQLLQRPS 400
324 DGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPAGSRAHSSHLKSKK 373
401 HLQPPSYGPVLSPMNKVHGGVNKLPSVNQLVGQPPPHSSAATPNLGPVGS 450
374 GQSTSRHKKFMFKTEGPDSD..... 393

FIG. 2

```
  1 TGCCTCCCCGCCCGCGCACCCGCCCCGAGGCCTGTGCTCCTGCGAAGGGG 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 TGCCTCCCCGCCCGCGCACCCGCCCCGAGGCCTGTGCTCCTGCGAAGGGG 50

 51 ACGCAGCGAAGCCGGGGCCCGCGCCAGGCCGGCCGGGACGGACGCCGATG 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 ACGCAGCGAAGCCGGGGCCCGCGCCAGGCCGGCCGGGACGGACGCCGATG 100

101 CCCGGAGCTGCGACGGCTGCAGAGCGAGCTGCCCTCGGAGGCCGGTGTGA 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 CCCGGAGCTGCGACGGCTGCAGAGCGAGCTGCCCTCGGAGGCCGGTGTGA 150

151 GGAAGATGGCCCAGTCCACCACCACCTCCCCCGATGGGGGCACCACGTTT 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 GGAAGATGGCCCAGTCCACCACCACCTCCCCCGATGGGGGCACCACGTTT 200

201 GAGCACCTCTGGAGCTCTCTGGAACCAGACAGCACCTACTTCGACCTTCC 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 GAGCACCTCTGGAGCTCTCTGGAACCAGACAGCACCTACTTCGACCTTCC 250

251 CCAGTCAAGCCGGGGGGAATAATGAGGTGGTGGGTGGCACGGATTCCAGCA 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 CCAGTCAAGCCGGGGGGAATAATGAGGTGGTGGGTGGCACGGATTCCAGCA 300

301 TGGACGTCTTCCACCTAGAGGGCATGACCACATCTGTCATGGCCCAGTTC 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 TGGACGTCTTCCACCTAGAGGGCATGACCACATCTGTCATGGCCCAGTTC 350

351 AATTTGCTGAGCAGCACCATGGACCAGATGAGCAGCCGCGCTGCCTCGGC 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 AATTTGCTGAGCAGCACCATGGACCAGATGAGCAGCCGCGCTGCCTCGGC 400

401 CAGCCCGTACACCCCGGAGCACGCCGCCAGCGTGCCCACCCATTCACCCT 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 CAGCCCGTACACCCCGGAGCACGCCGCCAGCGTGCCCACCCATTCACCCT 450

451 ACGCACAGCCCAGCTCCACCTTCGACACCATGTCGCCCGCGCCTGTCATC 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 ACGCACAGCCCAGCTCCACCTTCGACACCATGTCGCCCGCGCCTGTCATC 500

501 CCCTCCAACACCGACTATCCCGGACCCCACCACTTCGAGGTCACTTTCCA 550
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 CCCTCCAACACCGACTATCCCGGACCCCACCACTTCGAGGTCACTTTCCA 550

551 GCAGTCCAGCACGGCCAAGTCAGCCACCTGGACGTACTCCCCACTCTTGA 600
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 GCAGTCCAGCACGGCCAAGTCAGCCACCTGGACGTACTCCCCACTCTTGA 600

601 AGAAACTCTACTGCCAGATCGCCAAGACATGCCCCATCCAGATCAAGGTG 650
    ||||||||||||||||||||||||||||||||||||||||||||||||||
601 AGAAACTCTACTGCCAGATCGCCAAGACATGCCCCATCCAGATCAAGGTG 650

651 TCCGCCCCACCGCCCCCGGGCACCGCCATCCGGGCCATGCCTGTCTACAA 700
    ||||||||||||||||||||||||||||||||||||||||||||||||||
651 TCCGCCCCACCGCCCCCGGGCACCGCCATCCGGGCCATGCCTGTCTACAA 700

701 GAAGGCGGAGCACGTGACCGACATCGTGAAGCGCTGCCCCAACCACGAGC 750
    ||||||||||||||||||||||||||||||||||||||||||||||||||
701 GAAGGCGGAGCACGTGACCGACATCGTGAAGCGCTGCCCCAACCACGAGC 750

751 TCGGGAGGGACTTCAACGAAGGACAGTCTGCCCCAGCCAGCCACCTCATC 800
    ||||||||||||||||||||||||||||||||||||||||||||||||||
751 TCGGGAGGGACTTCAACGAAGGACAGTCTGCCCCAGCCAGCCACCTCATC 800

801 CGTGTGGAAGGCAATAATCTCTCGCAGTATGTGGACGACCCTGTCACCGG 850
    ||||||||||||||||||||||||||||||||||||||||||||||||||
801 CGTGTGGAAGGCAATAATCTCTCGCAGTATGTGGACGACCCTGTCACCGG 850

851 CAGGCAGAGCGTCGTGGTGCCCTATGAGCCACCACAGGTGGGGACAGAAT 900
    ||||||||||||||||||||||||||||||||||||||||||||||||||
851 CAGGCAGAGCGTCGTGGTGCCCTATGAGCCACCACAGGTGGGGACAGAAT 900
```

FIG. 3 cont.

FIG. 3A

```
 901 TCACCACCATCCTGTACAACTTCATGTGTAACAGCAGCTGTGTGGGGGGC 950
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 901 TCACCACCATCCTGTACAACTTCATGTGTAACAGCAGCTGTGTGGGGGGC 950

 951 ATGAACCGACGGCCCATCCTCATCATCATCACCCTGGAGACGCGGGATGG 1000
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 951 ATGAACCGACGGCCCATCCTCATCATCATCACCCTGGAGACGCGGGATGG 1000

1001 GCAGGTGCTGGGCCGCCGGTCCTTCGAGGGCCGCATCTGCGCCTGTCCTG 1050
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1001 GCAGGTGCTGGGCCGCCGGTCCTTCGAGGGCCGCATCTGCGCCTGTCCTG 1050

1051 GCCGCGACCGAAAAGCCGATGAGGACCACTACCGGGAGCAGCAGGCCTTG 1100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1051 GCCGCGACCGAAAAGCCGATGAGGACCACTACCGGGAGCAGCAGGCCTTG 1100

1101 AATGAGAGCTCCGCCAAGAACGGGGCTGCCAGCAAGCGCGCCTTCAAGCA 1150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1101 AATGAGAGCTCCGCCAAGAACGGGGCTGCCAGCAAGCGCGCCTTCAAGCA 1150

1151 GAGTCCCCCTGCCGTCCCCGCCCTGGGCCCGGGTGTGAAGAAGCGGCGGC 1200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1151 GAGTCCCCCTGCCGTCCCCGCCCTGGGCCCGGGTGTGAAGAAGCGGCGGC 1200

1201 ACGGAGACGAGGACACGTACTACCTGCAGGTGCGAGGCCGCGAGAACTTC 1250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1201 ACGGAGACGAGGACACGTACTACCTGCAGGTGCGAGGCCGCGAGAACTTC 1250

1251 GAGATCCTGATGAAGCTGAAGGAGAGCCTGGAGCTGATGGAGTTGGTGCC 1300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1251 GAGATCCTGATGAAGCTGAAGGAGAGCCTGGAGCTGATGGAGTTGGTGCC 1300

1301 GCAGCCGCTGGTAGACTCCTATCGGCAGCAGCAGCAGCTCCTACAGAGGC 1350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1301 GCAGCCGCTGGTAGACTCCTATCGGCAGCAGCAGCAGCTCCTACAGAGGC 1350

1351 CGAGTCACCTACAGCCCCCATCCTACGGGCCGGTCCTCTCGCCCATGAAC 1400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1351 CGAGTCACCTACAGCCCCCATCCTACGGGCCGGTCCTCTCGCCCATGAAC 1400

1401 AAGGTGCACGGGGGCGTGAACAAGCTGCCCTCCGTCAACCAGCTGGTGGG 1450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1401 AAGGTGCACGGGGGCGTGAACAAGCTGCCCTCCGTCAACCAGCTGGTGGG 1450

1451 CCAGCCTCCCCCCGCACAGCTCGGCAGCTACACCCAACCTGGGACCTGTGG 1500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1451 CCAGCCTCCCCCCGCACAGCTCGGCAGCTACACCCAACCTGGGACCTGTGG 1500

1501 GCTCTGGGATGCTCAACAACCACGGCCACGCAGTGCCAGCCAACAGCGAG 1550
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1501 GCTCTGGGATGCTCAACAACCACGGCCACGCAGTGCCAGCCAACAGCGAG 1550

1551 ATGACCAGCAGCCACGGCACCCAGTCCATGGTCTCGGGGTCCCACTGCAC 1600
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1551 ATGACCAGCAGCCACGGCACCCAGTCCATGGTCTCGGGGTCCCACTGCAC 1600

1601 TCCGCCACCCCCCTACCACGCCGACCCCAGCCTCGTCAGTTTTTTAACAG 1650
     ||||||||||||||||||||||||||||||||||||
1601 TCCGCCACCCCCCTACCACGCCGACCCCAGCCTCGTC..............  1637


1701 AGCATTTACCACCTGCAGAACCTGACCATCGAGGACCTGGGGGCCCTGAA 1750
     ||||||||||||||||||||
1638 ....................AGGACCTGGGGGCCCTGAA 1656

1751 GATCCCCGAGCAGTATCGCATGACCATCTGGCGGGGCCTGCAGGACCTGA 1800
     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

FIG. 3 cont.

FIG. 3B

```
1657 GATCCCCGAGCAGTATCGCATGACCATCTGGCGGGGCCTGCAGGACCTGA 1706

1801 AGCAGGGCCACGACTACGGCGCCGCCGCGCAGCAGCTGCTCCGCTCCAGC 1850
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1707 AGCAGGGCCACGACTACGGCGCCGCCGCGCAGCAGCTGCTCCGCTCCAGC 1756

1851 AACGCGGCCGCCATTTCCATCGGCGGCTCCGGGGAGCTGCAGCGCCAGCG 1900
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1757 AACGCGGCCGCCATTTCCATCGGCGGCTCCGGGGAGCTGCAGCGCCAGCG 1806

1901 GGTCATGGAGGCCGTGCACTTCCGCGTGCGCCACACCATCACCATCCCCA 1950
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1807 GGTCATGGAGGCCGTGCACTTCCGCGTGCGCCACACCATCACCATCCCCA 1856

1951 ACCGCGGCGGCCCCCGGCGCCGGCCCCGACGAGTGGGCGGACTTCGGCTTC 2000
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1857 ACCGCGGCGGCCCCCGGCGCCGGCCCCGACGAGTGGGCGGACTTCGGCTTC 1906

2001 GACCTGCCCGACTGCAAGGCCCGCAAGCAGCCCATCAAGGAGGAGTTCAC 2050
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1907 GACCTGCCCGACTGCAAGGCCCGCAAGCAGCCCATCAAGGAGGAGTTCAC 1956

2051 GGAGGCCGAGATCCACTGAGGGGCCGGGCCCAGCCAGAGCCTGTGCCACC 2100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1957 GGAGGCCGAGATCCACTGAGGGGCCGGGCCCAGCCAGAGCCTGTGCCACC 2006

2101 GCCCAGAGACCCAGGCCGCCTCGCTCTC 2128
     ||||||||||||||||||||||||||||
2007 GCCCAGAGACCCAGGCCGCCTCGCTCTC 2034
```

FIG. 3C

~~FIG. 3 cont.~~

```
     1  TGCCTCCCCGCCCGCGCACCCGCCCCGAGGCCTGTGCTCCTGCGAAGGGGACGCAGCGAA   60
    61  GCCGGGGCCCGCGCCAGGCCGGCCGGGACGGACGCCGATGCCCGGAGCTGCGACGGCTGC  120
   121  AGAGCGAGCTGCCCTCGGAGGCCGGTGTGAGGAAGATGGCCCAGTCCACCACCACCTCCC  180
   -10                                   M  A  Q  S  T  T  T  S  P    9
   181  CCGATGGGGGCACCACGTTTGAGCACCTCTGGAGCTCTCTGGAACCAGACAGCACCTACT  240
    10    D  G  G  T  T  F  E  H  L  W  S  S  L  E  P  D  S  T  Y  F   29
   241  TCGACCTTCCCCAGTCAAGCCGGGGGAATAATGAGGTGGTGGGTGGCACGGATTCCAGCA  300
    30    D  L  P  Q  S  S  R  G  N  N  E  V  V  G  G  T  D  S  S  M   49
   301  TGGACGTCTTCCACCTAGAGGGCATGACCACATCTGTCATGGCCCAGTTCAATTTGCTGA  360
    50    D  V  F  H  L  E  G  M  T  T  S  V  M  A  Q  F  N  L  L  S   69
   361  GCAGCACCATGGACCAGATGAGCAGCCGCGCTGCCTCGGCCAGCCCGTACACCCCGGAGC  420
    70    S  T  M  D  Q  M  S  S  R  A  A  S  A  S  P  Y  T  P  E  H   89
   421  ACGCCGCCAGCGTGCCCACCCATTCACCCTACGCACAGCCCAGCTCCACCTTCGACACCA  480
    90    A  A  S  V  P  T  H  S  P  Y  A  Q  P  S  S  T  F  D  T  M  109
   481  TGTCGCCCGCGCCTGTCATCCCCTCCAACACCGACTATCCCGGACCCCACCACTTCGAGG  540
   110    S  P  A  P  V  I  P  S  N  T  D  Y  P  G  P  H  H  F  E  V  129
   541  TCACTTTCCAGCAGTCCAGCACGGCCAAGTCAGCCACCTGGACGTACTCCCCACTCTTGA  600
   130    T  F  Q  Q  S  S  T  A  K  S  A  T  W  T  Y  S  P  L  L  K  149
   601  AGAAACTCTACTGCCAGATCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCGCCCCAC  660
   150    K  L  Y  C  Q  I  A  K  T  C  P  I  Q  I  K  V  S  A  P  P  169
   661  CGCCCCCGGGCACCGCCATCCGGGCCATGCCTGTCTACAAGAAGGCGGAGCACGTGACCG  720
   170    P  P  G  T  A  I  R  A  M  P  V  Y  K  K  A  E  H  V  T  D  189
   721  ACATCGTGAAGCGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACGAAGGACAGTCTG  780
   190    I  V  K  R  C  P  N  H  E  L  G  R  D  F  N  E  G  Q  S  A  209
   781  CCCCAGCCAGCCACCTCATCCGTGTGGAAGGCAATAATCTCTCGCAGTATGTGGACGACC  840
   210    P  A  S  H  L  I  R  V  E  G  N  N  L  S  Q  Y  V  D  D  P  229
   841  CTGTCACCGGCAGGCAGAGCGTCGTGGTGCCCTATGAGCCACCACAGGTGGGGACAGAAT  900
   230    V  T  G  R  Q  S  V  V  V  P  Y  E  P  P  Q  V  G  T  E  F  249
   901  TCACCACCATCCTGTACAACTTCATGTGTAACAGCAGCTGTGTGGGGGGCATGAACCGAC  960
   250    T  T  I  L  Y  N  F  M  C  N  S  S  C  V  G  G  M  N  R  R  269
   961  GGCCCATCCTCATCATCATCACCCTGGAGACGCGGGATGGGCAGGTGCTGGGCCGCCGGT 1020
   270    P  I  L  I  I  I  T  L  E  T  R  D  G  Q  V  L  G  R  R  S  289
  1021  CCTTCGAGGGCCGCATCTGCGCCTGTCCTGGCCGCGACCGAAAAGCCGATGAGGACCACT 1080
   290    F  E  G  R  I  C  A  C  P  G  R  D  R  K  A  D  E  D  H  Y  309
  1081  ACCGGGAGCAGCAGGCCTTGAATGAGAGCTCCGCCAAGAACGGGGCTGCCAGCAAGCGCG 1140
   310    R  E  Q  Q  A  L  N  E  S  S  A  K  N  G  A  A  S  K  R  A  329
  1141  CCTTCAAGCAGAGTCCCCCTGCCGTCCCCGCCCTGGGCCCGGGTGTGAAGAAGCGGCGGC 1200
   330    F  K  Q  S  P  P  A  V  P  A  L  G  P  G  V  K  K  R  R  H  349
  1201  ACGGAGACGAGGACACGTACTACCTGCAGGTGCGAGGCCGCGAGAACTTCGAGATCCTGA 1260
   350    G  D  E  D  T  Y  Y  L  Q  V  R  G  R  E  N  F  E  I  L  M  369
  1261  TGAAGCTGAAGGAGAGCCTGGAGCTGATGGAGTTGGTGCCGCAGCCGCTGGTAGACTCCT 1320
   370    K  L  K  E  S  L  E  L  M  E  L  V  P  Q  P  L  V  D  S  Y  389
  1321  ATCGGCAGCAGCAGCAGCTCCTACAGAGGCCGAGTCACCTACAGCCCCCATCCTACGGGC 1380
   390    R  Q  Q  Q  Q  L  L  Q  R  P  S  H  L  Q  P  P  S  Y  G  P  409
  1381  CGGTCCTCTCGCCCATGAACAAGGTGCACGGGGGCGTGAACAAGCTGCCCTCCGTCAACC 1440
   410    V  L  S  P  M  N  K  V  H  G  G  V  N  K  L  P  S  V  N  Q  429
  1441  AGCTGGTGGGCCAGCCTCCCCCGCACAGCTCGGCAGCTACACCCAACCTGGGACCTGTGG 1500
   430    L  V  G  Q  P  P  P  H  S  S  A  A  T  P  N  L  G  P  V  G  449
  1501  GCTCTGGGATGCTCAACAACCACGGCCACGCAGTGCCAGCCAACAGCGAGATGACCAGCA 1560
   450    S  G  M  L  N  N  H  G  H  A  V  P  A  N  S  E  M  T  S  S  469
  1561  GCCACGGCACCCAGTCCATGGTCTCGGGGTCCCACTGCACTCCGCCACCCCCCTACCACG 1620
   470    H  G  T  Q  S  M  V  S  G  S  H  C  T  P  P  P  P  Y  H  A  489
  1621  CCGACCCCAGCCTCGTCAGTTTTTTAACAGGATTGGGGTGTCCAAACTGCATCGAGTATT 1680
   490    D  P  S  L  V  S  F  L  T  G  L  G  C  P  N  C  I  E  Y  P  509
```

FIG. 4

FIG. 4A

```
1681  TCACGTCCCAGGGGTTACAGAGCATTTACCACCTGCAGAACCTGACCATCGAGGACCTGG  1740
 510    T  S  Q  G  L  Q  S  I  Y  H  L  Q  N  L  T  I  E  D  L  G   529
1741  GGGCCCTGAAGATCCCCGAGCAGTATCGCATGACCATCTGGCGGGGCCTGCAGGACCTGA  1800
 530    A  L  K  I  P  E  Q  Y  R  M  T  I  W  R  G  L  Q  D  L  K   549
1801  AGCAGGGCCACGACTACGGCGCCGCCGCGCAGCAGCTGCTCCGCTCCAGCAACGCGGCCG  1860
 550    Q  G  H  D  Y  G  A  A  A  Q  Q  L  L  R  S  S  N  A  A  A   569
1861  CCATTTCCATCGGCGGCTCCGGGGAGCTGCAGCGCCAGCGGGTCATGGAGGCCGTGCACT  1920
 570    I  S  I  G  G  S  G  E  L  Q  R  Q  R  V  M  E  A  V  H  F   589
1921  TCCGCGTGCGCCACACCATCACCATCCCCAACCGCGGCGGCCCCGGCGCCGGCCCCGACG  1980
 590    R  V  R  H  T  I  T  I  P  N  R  G  G  P  G  A  G  P  D  E   609
1981  AGTGGGCGGACTTCGGCTTCGACCTGCCCGACTGCAAGGCCCGCAAGCAGCCCATCAAGG  2040
 610    W  A  D  F  G  F  D  L  P  D  C  K  A  R  K  Q  P  I  K  E   629
2041  AGGAGTTCACGGAGGCCGAGATCCACTGAGGGGCCGGGCCCAGCCAGAGCCTGTGCCACC  2100
 630    E  F  T  E  A  E  I  H  *                                    649
2101  GCCCAGAGACCCAGGCCGCCTCGCTCTCCTTCCTGTGTCCAAAACTGCCTCCGGAGGCAG  2160
2161  GGCCTCCAGGCTGTGCCCGGGGAAAGGCAAGGTCCGGCCCATGCCCCGGCACCTCACCGG  2220
2221  CCCCAGGAGAGGCCCAGCCACCAAAGCCGCCTGCGGACAGCCTGAGTCACCTGCAGAACC  2280
2281  TTCTGGAGCTGCCCTAATGCTGGGCTTGCGGGGCAGGGGCCGGCCCACTCTCAGCCCTGC  2340
2341  CACTGCCGGGCGTGCTCCATGGCAGGCGTGGGTGGGGACCGCAGTGTCAGCTCCGACCTC  2400
2401  CAGGCCTCATCCTAGAGACTCTGTCATCTGCCGATCAAGCAAGGTCCTTCCAGAGGAAAG  2460
2461  AATCCTCTTCGCTGGTGGACTGCCAAAAAGTATTTTGCGACATCTTTTGGTTCTGGAGAG  2520
2521  TGGTGAGCAGCCAAGCGACTGTGTCTGAAACACCGTGCATTTTCAGGGAATGTCCCTAAC  2580
2581  GGGCTGGGGACTCTCTCTGCTGGACTTGGGAGTGGCCTTTGCCCCCAGCACACTGTATTC  2640
2641  TGCGGGACCGCCTCCTTCCTGCCCCTAACAACCACCAAAGTGTTGCTGAAATTGGAGAAA  2700
2701  ACTGGGGAAGGCGCAACCCCTCCCAGGTGCGGGAAGCATCTGGTACCGCCTCGGCCAGTG  2760
2761  CCCCTCAGCCTGGCCACAGTCACCTCTCCTTGGGGAACCCTGGGCAGAAAGGGACAGCCT  2820
2821  GTCCTTAGAGGACCGGAAATTGTCAATATTTGATAAAATGATACCCTTTTCTAC  2874
```

FIG. 4B

~~FIG. 4 cont.~~

```
     1  TGCCTCCCCGCCCGCGCACCCGCCCCGAGGCCTGTGCTCCTGCGAAGGGGACGCAGCGAA   60
    61  GCCGGGGCCCCGCGCCAGGCCGGCCGGGACGGACGCCGATGCCCCGGAGCTGCGACGGCTGC  120
   121  AGAGCGAGCTGCCCTCGGAGGCCGGTGTGAGGAAGATGGCCCAGTCCACCACCACCTCCC    180
   -10                                    M  A  Q  S  T  T  T  S  P      9
   181  CCGATGGGGGCACCACGTTTGAGCACCTCTGGAGCTCTCTGGAACCAGACAGCACCTACT    240
    10    D  G  G  T  T  F  E  H  L  W  S  S  L  E  P  D  S  T  Y  F     29
   241  TCGACCTTCCCCAGTCAAGCCGGGGGAATAATGAGGTGGTGGGTGGCACGGATTCCAGCA    300
    30    D  L  P  Q  S  S  R  G  N  N  E  V  V  G  G  T  D  S  S  M     49
   301  TGGACGTCTTCCACCTAGAGGGCATGACCACATCTGTCATGGCCCAGTTCAATTTGCTGA    360
    50    D  V  F  H  L  E  G  M  T  T  S  V  M  A  Q  F  N  L  L  S     69
   361  GCAGCACCATGGACCAGATGAGCAGCCGCGCTGCCTCGGCCAGCCCGTACACCCCGGAGC    420
    70    S  T  M  D  Q  M  S  S  R  A  A  S  A  S  P  Y  T  P  E  H     89
   421  ACGCCGCCAGCGTGCCCACCCATTCACCCTACGCACAGCCCAGCTCCACCTTCGACACCA    480
    90    A  A  S  V  P  T  H  S  P  Y  A  Q  P  S  S  T  F  D  T  M    109
   481  TGTCGCCCGCGCCTGTCATCCCCTCCAACACCGACTATCCCGGACCCCACCACTTCGAGG    540
   110    S  P  A  P  V  I  P  S  N  T  D  Y  P  G  P  H  H  F  E  V    129
   541  TCACTTTCCAGCAGTCCAGCACGGCCAAGTCAGCCACCTGGACGTACTCCCCACTCTTGA    600
   130    T  F  Q  Q  S  S  T  A  K  S  A  T  W  T  Y  S  P  L  L  K    149
   601  AGAAACTCTACTGCCAGATCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCGCCCCAC    660
   150    K  L  Y  C  Q  I  A  K  T  C  P  I  Q  I  K  V  S  A  P  P    169
   661  CGCCCCCGGGCACCGCCATCCGGGCCATGCCTGTCTACAAGAAGGCGGAGCACGTGACCG    720
   170    P  P  G  T  A  I  R  A  M  P  V  Y  K  K  A  E  H  V  T  D    189
   721  ACATCGTGAAGCGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACGAAGGACAGTCTG    780
   190    I  V  K  R  C  P  N  H  E  L  G  R  D  F  N  E  G  Q  S  A    209
   781  CCCCAGCCAGCCACCTCATCCGTGTGGAAGGCAATAATCTCTCGCAGTATGTGGACGACC    840
   210    P  A  S  H  L  I  R  V  E  G  N  N  L  S  Q  Y  V  D  D  P    229
   841  CTGTCACCGGCAGGCAGAGCGTCGTGGTGCCCTATGAGCCACCACAGGTGGGGACAGAAT    900
   230    V  T  G  R  Q  S  V  V  V  P  Y  E  P  P  Q  V  G  T  E  F    249
   901  TCACCACCATCCTGTACAACTTCATGTGTAACAGCAGCTGTGTGGGGGGCATGAACCGAC    960
   250    T  T  I  L  Y  N  F  M  C  N  S  S  C  V  G  G  M  N  R  R    269
   961  GGCCCATCCTCATCATCATCACCCTGGAGACGCGGGATGGGCAGGTGCTGGGCCGCCGGT   1020
   270    P  I  L  I  I  I  T  L  E  T  R  D  G  Q  V  L  G  R  R  S    289
  1021  CCTTCGAGGGCCGCATCTGCGCCTGTCCTGGCCGCGACCGAAAAGCCGATGAGGACCACT   1080
   290    F  E  G  R  I  C  A  C  P  G  R  D  R  K  A  D  E  D  H  Y    309
  1081  ACCGGGAGCAGCAGGCCTTGAATGAGAGCTCCGCCAAGAACGGGGCTGCCAGCAAGCGCG   1140
   310    R  E  Q  Q  A  L  N  E  S  S  A  K  N  G  A  A  S  K  R  A    329
  1141  CCTTCAAGCAGAGTCCCCCTGCCGTCCCCGCCCTGGGCCCGGGTGTGAAGAAGCGGCGGC   1200
   330    F  K  Q  S  P  P  A  V  P  A  L  G  P  G  V  K  K  R  R  H    349
  1201  ACGGAGACGAGGACACGTACTACCTGCAGGTGCGAGGCCGCGAGAACTTCGAGATCCTGA   1260
   350    G  D  E  D  T  Y  Y  L  Q  V  R  G  R  E  N  F  E  I  L  M    369
  1261  TGAAGCTGAAGGAGAGCCTGGAGCTGATGGAGTTGGTGCCGCAGCCGCTGGTAGACTCCT   1320
   370    K  L  K  E  S  L  E  L  M  E  L  V  P  Q  P  L  V  D  S  Y    389
  1321  ATCGGCAGCAGCAGCAGCTCCTACAGAGGCCGAGTCACCTACAGCCCCCATCCTACGGGC   1380
   390    R  Q  Q  Q  Q  L  L  Q  R  P  S  H  L  Q  P  P  S  Y  G  P    409
  1381  CGGTCCTCTCGCCCATGAACAAGGTGCACGGGGGCGTGAACAAGCTGCCCTCCGTCAACC   1440
   410    V  L  S  P  M  N  K  V  H  G  G  V  N  K  L  P  S  V  N  Q    429
  1441  AGCTGGTGGGCCAGCCTCCCCCGCACAGCTCGGCAGCTACACCCAACCTGGGACCTGTGG   1500
   430    L  V  G  Q  P  P  P  H  S  S  A  A  T  P  N  L  G  P  V  G    449
  1501  GCTCTGGGATGCTCAACAACCACGGCCACGCAGTGCCAGCCAACAGCGAGATGACCAGCA   1560
   450    S  G  M  L  N  N  H  G  H  A  V  P  A  N  S  E  M  T  S  S    469
  1561  GCCACGGCACCCAGTCCATGGTCTCGGGGTCCCACTGCACTCCGCCACCCCCCTACCACG   1620
   470    H  G  T  Q  S  M  V  S  G  S  H  C  T  P  P  P  P  Y  H  A    489
  1621  CCGACCCCAGCCTCGTCAGGACCTGGGGGCCCTGAAGATCCCCGAGCAGTATCGCATGAC   1680
   490    D  P  S  L  V  R  T  W  G  P  *                              509
  1681  CATCTGGCGGGGCCTGCAGGACCTGAAGCAGGGCCACGACTACGGCGCCGCCGCGCAGCA   1740
  1741  GCTGCTCCGCTCCAGCAACGCGGCCGCCATTTCCATCGGCGGCTCCGGGGAGCTGCAGCG   1800
  1801  CCAGCGGGTCATGGAGGCCGTGCACTTCCGCGTGCGCCACACCATCACCATCCCCAACCG   1860
  1861  CGGCGGCCCCGGCGCCGGCCCCGACGAGTGGGCGGACTTCGGCTTCGACCTGCCCGACTG   1920
  1921  CAAGGCCCGCAAGCAGCCCATCAAGGAGGAGTTCACGGAGGCCGAGATCCACTGAGGGGC   1980
  1981  CGGGCCCAGCCAGAGCCTGTGCCACCGCCCAGAGACCCAGGCCGCCTCGCTCTC  2034
```

FIG.5

```
   1  GCGAGCTGCCCTCGGAGGCCGGCGTGGGGAAGATGGCCCAGTCCACCGCCACCTCCCCTG   60
  -9                                  M  A  Q  S  T  A  T  S  P  D   10
  61  ATGGGGGCACCACGTTTGAGCACCTCTGGAGCTCTCTGGAACCAGACAGCACCTACTTCG  120
  11    G  G  T  T  F  E  H  L  W  S  S  L  E  P  D  S  T  Y  F  D   30
 121  ACCTTCCCCAGTCAAGCCGGGGGAATAATGAGGTGGTGGGCGGAACGGATTCCAGCATGG  180
  31    L  P  Q  S  S  R  G  N  N  E  V  V  G  G  T  D  S  S  M  D   50
 181  ACGTCTTCCACCTGGAGGGCATGACTACATCTGTCATGGCCCAGTTCAATCTGCTGAGCA  240
  51    V  F  H  L  E  G  M  T  T  S  V  M  A  Q  F  N  L  L  S  S   70
 241  GCACCATGGACCAGATGAGCAGCCGCGCGGCCTCGGCCAGCCCCTACACCCCAGAGCACG  300
  71    T  M  D  Q  M  S  S  R  A  A  S  A  S  P  Y  T  P  E  H  A   90
 301  CCGCCAGCGTGCCCACCCACTCGCCCTACGCACAACCCAGCTCCACCTTCGACACCATGT  360
  91    A  S  V  P  T  H  S  P  Y  A  Q  P  S  S  T  F  D  T  M  S  110
 361  CGCCGGCGCCTGTCATCCCCTCCAACACCGACTACCCCGGACCCCACCACTTTGAGGTCA  420
 111    P  A  P  V  I  P  S  N  T  D  Y  P  G  P  H  H  F  E  V  T  130
 421  CTTTCCAGCAGTCCAGCACGGCCAAGTCAGCCACCTGGACGTACTCCCCGCTCTTGAAGA  480
 131    F  Q  Q  S  S  T  A  K  S  A  T  W  T  Y  S  P  L  L  K  K  150
 481  AACTCTACTGCCAGATCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCACCCCGCCAC  540
 151    L  Y  C  Q  I  A  K  T  C  P  I  Q  I  K  V  S  T  P  P  P  170
 541  CCCCAGGCACTGCCATCCGGGCCATGCCTGTTTACAAGAAAGCGGAGCACGTGACCGACG  600
 171    P  G  T  A  I  R  A  M  P  V  Y  K  K  A  E  H  V  T  D  V  190
 601  TCGTGAAACGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACGAAGGACAGTCTGCTC  660
 191    V  K  R  C  P  N  H  E  L  G  R  D  F  N  E  G  Q  S  A  P  210
 661  CAGCCAGCCACCTCATCCGCGTGGAAGGCAATAATCTCTCGCAGTATGTGGATGACCCTG  720
 211    A  S  H  L  I  R  V  E  G  N  N  L  S  Q  Y  V  D  D  P  V  230
 721  TCACCGGCAGGCAGAGCGTCGTGGTGCCCTATGAGCCACCACAGGTGGGGACGGAATTCA  780
 231    T  G  R  Q  S  V  V  V  P  Y  E  P  P  Q  V  G  T  E  F  T  250
 781  CCACCATCCTGTACAACTTCATGTGTAACAGCAGCTGTGTAGGGGGCATGAACCGGCGGC  840
 251    T  I  L  Y  N  F  M  C  N  S  S  C  V  G  G  M  N  R  R  P  270
 841  CCATCCTCATCATCATCACCCTGGAGATGCGGGATGGGCAGGTGCTGGGCCGCCGGTCCT  900
 271    I  L  I  I  I  T  L  E  M  R  D  G  Q  V  L  G  R  R  S  F  290
 901  TTGAGGGCCGCATCTGCGCCTGTCCTGGCCGCGACCGAAAAGCTGATGAGGACCACTACC  960
 291    E  G  R  I  C  A  C  P  G  R  D  R  K  A  D  E  D  H  Y  R  310
 961  GGGAGCAGCAGGCCCTGAACGAGAGCTCCGCCAAGAACGGGGCCGCCAGCAAGCGTGCCT 1020
 311    E  Q  Q  A  L  N  E  S  S  A  K  N  G  A  A  S  K  R  A  F  330
1021  TCAAGCAGAGCCCCCCTGCCGTCCCCGCCCTTGGTGCCGGTGTGAAGAAGCGGCGGCATG 1080
 331    K  Q  S  P  P  A  V  P  A  L  G  A  G  V  K  K  R  R  H  G  350
1081  GAGACGAGGACACGTACTACCTTCAGGTGCGAGGCCGGGAGAACTTTGAGATCCTGATGA 1140
 351    D  E  D  T  Y  Y  L  Q  V  R  G  R  E  N  F  E  I  L  M  K  370
1141  AGCTGAAAGAGAGCCTGGAGCTGATGGAGTTGGTGCCGCAGCCACTGGTGGACTCCTATC 1200
 371    L  K  E  S  L  E  L  M  E  L  V  P  Q  P  L  V  D  S  Y  R  390
1201  GGCAGCAGCAGCAGCTCCTACAGAGGCCGAGTCACCTACAGCCCCCGTCCTACGGGCCGG 1260
 391    Q  Q  Q  Q  L  L  Q  R  P  S  H  L  Q  P  P  S  Y  G  P  V  410
1261  TCCTCTCGCCCATGAACAAGGTGCACGGGGGCATGAACAAGCTGCCCTCCGTCAACCAGC 1320
 411    L  S  P  M  N  K  V  H  G  G  M  N  K  L  P  S  V  N  Q  L  430
1321  TGGTGGGCCAGCCTCCCCCGCACAGTTCGGCAGCTACACCCAACCTGGGGCCCGTGGGCC 1380
 431    V  G  Q  P  P  P  H  S  S  A  A  T  P  N  L  G  P  V  G  P  450
1381  CCGGGATGCTCAACAACCATGGCCACGCAGTGCCAGCCAACGGCGAGATGAGCAGCAGCC 1440
 451    G  M  L  N  N  H  G  H  A  V  P  A  N  G  E  M  S  S  S  H  470
```

FIG. 6A

~~FIG. 6~~

```
1441  ACAGCGCCCAGTCCATGGTCTCGGGGTCCCACTGCACTCCGCCACCCCCCTACCACGCCG  1500
 471    S  A  Q  S  M  V  S  G  S  H  C  T  P  P  P  P  Y  H  A  D   490
1501  ACCCCAGCCTCGTCAGTTTTTTAACAGGATTGGGGTGTCCAAACTGCATCGAGTATTTCA  1560
 491    P  S  L  V  S  F  L  T  G  L  G  C  P  N  C  I  E  Y  F  T   510
1561  CCTCCCAAGGGTTACAGAGCATTTACCACCTGCAGAACCTGACCATTGAGGACCTGGGGG  1620
 511    S  Q  G  L  Q  S  I  Y  H  L  Q  N  L  T  I  E  D  L  G  A   530
1621  CCCTGAAGATCCCCGAGCAGTACCGCATGACCATCTGGCGGGGCCTGCAGGACCTGAAGC  1680
 531    L  K  I  P  E  Q  Y  R  M  T  I  W  R  G  L  Q  D  L  K  Q   550
1681  AGGGCCACGACTACAGCACCGCGCAGCAGCTGCTCCGCTCTAGCAACGCGGCCACCATCT  1740
 551    G  H  D  Y  S  T  A  Q  Q  L  L  R  S  S  N  A  A  T  I  S   570
1741  CCATCGGCGGCTCAGGGGAACTGCAGCGCCAGCGGGTCATGGAGGCCGTGCACTTCCGCG  1800
 571    I  G  G  S  G  E  L  Q  R  Q  R  V  M  E  A  V  H  F  R  V   590
1801  TGCGCCACACCATCACCATCCCCAACCGCGGCGGCCCAGGCGGCGGCCCTGACGAGTGGG  1860
 591    R  H  T  I  T  I  P  N  R  G  G  P  G  G  G  P  D  E  W  A   610
1861  CGGACTTCGGCTTCGACCTGCCCGACTGCAAGGCCCGCAAGCAGCCCATCAAGGAGGAGT  1920
 611    D  F  G  F  D  L  P  D  C  K  A  R  K  Q  P  I  K  E  E  F   630
1921  TCACGGAGGCCGAGATCCACTGAGGGCCTCGCCTGGCTGCAGCCTGCGCCACCGCCCAGA  1980
 631    T  E  A  E  I  H  *                                          650
1981  GACCCAAGCTGCCTCCCCTCTCCTTCCTGTGTGTCCAAAACTGCCTCAGGAGGCAGGACC  2040
2041  TTCGGGCTGTGCCCGGGGAAAGGCAAGGTCCGGCCCATCCCCAGGCACCTCACAGGCCCC  2100
2101  AGGAAAGGCCCAGCCACCGAAGCCGCCTGTGGACAGCCTGAGTCACCTGCAGAACC  2156
```

FIG. 6B

~~FIG. 6 cont.~~

```
     1  TGATCTCCCTGTGGCCTGCAGGGGACTGAGCCAGGGAGTAGATGCCCTGAGACCCCAAGG   60
    61  GACACCCAAGGAAACCTTGCTGGCTTTGAGAAAGGGATCGTCTCTCTCCTGCCCAAGAGA  120
   121  AGCATGTGTATGGGCCCTGTGTATGAATCCTTGGGGCAGGCCCAGTTCAATTTGCTCAGC  180
     0     M  C  M  G  P  V  Y  E  S  L  G  Q  A  Q  F  N  L  L  S      19
   181  AGTGCCATGGACCAGATGGGCAGCCGTGCGGCCCCGGCGAGCCCCTACACCCCGGAGCAC  240
    20  S  A  M  D  Q  M  G  S  R  A  A  P  A  S  P  Y  T  P  E  H      39
   241  GCCGCCAGCGCGCCCACCCACTCGCCCTACGCGCAGCCCAGCTCCACCTTCGACACCATG  300
    40  A  A  S  A  P  T  H  S  P  Y  A  Q  P  S  S  T  F  D  T  M      59
   301  TCTCCGGCGCCTGTCATCCCTTCCAATACCGACTACCCCGGCCCCCACCACTTCGAGGTC  360
    60  S  P  A  P  V  I  P  S  N  T  D  Y  P  G  P  H  H  F  E  V      79
   361  ACCTTCCAGCAGTCGAGCACTGCCAAGTCGGCCACCTGGACATACTCCCCACTCTTGAAG  420
    80  T  F  Q  Q  S  S  T  A  K  S  A  T  W  T  Y  S  P  L  L  K      99
   421  AAGTTGTACTGTCAGATTGCTAAGACATGCCCCATCCAGATCAAAGTGTCCACACCACCA  480
   100  K  L  Y  C  Q  I  A  K  T  C  P  I  Q  I  K  V  S  T  P  P     119
   481  CCCCCGGGCACGGCCATCCGGGCCATGCCTGTCTACAAGAAGGCAGAGCATGTGACCGAC  540
   120  P  P  G  T  A  I  R  A  M  P  V  Y  K  K  A  E  H  V  T  D     139
   541  ATTGTTAAGCGCTGCCCCAACCACGAGCTTGGAAGGGACTTCAATGAAGGACAGTCTGCC  600
   140  I  V  K  R  C  P  N  H  E  L  G  R  D  F  N  E  G  Q  S  A     159
   601  CCGGCTAGCCACCTCATCCGTGTAGAAGGCAACAACCTCGCCCAGTACGTGGATGACCCT  660
   160  P  A  S  H  L  I  R  V  E  G  N  N  L  A  Q  Y  V  D  D  P     179
   661  GTCACCGGAAGGCAGAGTGTGGTTGTGCCGTATGAACCCCCACAGGTGGGAACAGAATTT  720
   180  V  T  G  R  Q  S  V  V  V  P  Y  E  P  P  Q  V  G  T  E  F     199
   721  ACCACCATCCTGTACAACTTCATGTGTAACAGCAGCTGTGTGGGGGGCATGAATCGGAGG  780
   200  T  T  I  L  Y  N  F  M  C  N  S  S  C  V  G  G  M  N  R  R     219
   781  CCCATCCTTGTCATCATCACCCTGGAGACCCGGGATGGACAGGTCCTGGGCCGCCGGTCT  840
   220  P  I  L  V  I  I  T  L  E  T  R  D  G  Q  V  L  G  R  R  S     239
   841  TTCGAGGGTCGCATCTGTGCCTGTCCTGGCCGTGACCGCAAAGCTGATGAAGACCATTAC  900
   240  F  E  G  R  I  C  A  C  P  G  R  D  R  K  A  D  E  D  H  Y     259
   901  CGGGAGCAACAGGCTCTGAATGAAAGTACCACCAAAAATGGAGCTGCCAGCAAACGTGCA  960
   260  R  E  Q  Q  A  L  N  E  S  T  T  K  N  G  A  A  S  K  R  A     279
   961  TTCAAGCAGAGCCCCCCTGCCATCCCTGCCCTGGGTACCAACGTGAAGAAGAGACGCCAC 1020
   280  F  K  Q  S  P  P  A  I  P  A  L  G  T  N  V  K  K  R  R  H     299
  1021  GGGGACGAGGACATGTTCTACATGCACGTGCGAGGCCGGGAGAACTTTGAGATCTTGATG 1080
   300  G  D  E  D  M  F  Y  M  H  V  R  G  R  E  N  F  E  I  L  M     319
  1081  AAAGTCAAGGAGAGCCTAGAACTGATGGAGCTTGTGCCCCAGCCTTTGGTTGACTCCTAT 1140
   320  K  V  K  E  S  L  E  L  M  E  L  V  P  Q  P  L  V  D  S  Y     339
  1141  CGACAGCAGCAGCAGCAGCAGCTCCTACAGAGGCCGAGTCACCTGCAGCCTCCATCCTAT 1200
   340  R  Q  Q  Q  Q  Q  Q  L  L  Q  R  P  S  H  L  Q  P  P  S  Y     359
  1201  GGGCCCGTGCTCTCCCCAATGAACAAGGTACACGGTGGTGTCAACAAACTGCCCTCCGTC 1260
   360  G  P  V  L  S  P  M  N  K  V  H  G  G  V  N  K  L  P  S  V     379
  1261  AACCAGCTGGTGGGCCAGCCTCCCCCGCACAGCTCAGCAGCTGGGCCCAACCTGGGGCCC 1320
   380  N  Q  L  V  G  Q  P  P  P  H  S  S  A  A  G  P  N  L  G  P     399
  1321  ATGGGCTCCGGGATGCTCAACAGCCACGGCCACAGCATGCCGGCCAATGGTGAGATGAAT 1380
   400  M  G  S  G  M  L  N  S  H  G  H  S  M  P  A  N  G  E  M  N     419
  1381  GGAGGCCACAGCTCCCAGACCATGGTTTCGGGATCCCACTGCACCCCGCCACCCCCCTAT 1440
   420  G  G  H  S  S  Q  T  M  V  S  G  S  H  C  T  P  P  P  P  Y     439
  1441  CATGCAGACCCCAGCCTCGTCAGTTTTTTGACAGGGTTGGGGTGTCCAAACTGCATCGAG 1500
   440  H  A  D  P  S  L  V  S  F  L  T  G  L  G  C  P  N  C  I  E     459
  1501  TGCTTCACTTCCCAAGGGTTGCAGAGCATCTACCACCTGCAGAACCTTACCATCGAGGAC 1560
   460  C  F  T  S  Q  G  L  Q  S  I  Y  H  L  Q  N  L  T  I  E  D     479
  1561  CTTGGGGCTCTGAAGGTCCCTGACCAGTACCGTATGACCATCTGGAGGGGCCTACAGGAC 1620
   480  L  G  A  L  K  V  P  D  Q  Y  R  M  T  I  W  R  G  L  Q  D     499
  1621  CTGAAGCAGAGCCATGACTGCGGCCAGCAACTGCTACGCTCCAGCAGCAACGCGGCCACC 1680
   500  L  K  Q  S  H  D  C  G  Q  Q  L  L  R  S  S  S  N  A  A  T     519
  1681  ATCTCCATCGGCGGCTCTGGCGAGCTGCAGCGGCAGCGGGTCATGGAAGCCGTGCATTTC 1740
   520  I  S  I  G  G  S  G  E  L  Q  R  Q  R  V  M  E  A  V  H  F     539
  1741  CGTGTGCGCCACACCATCACAATCCCCAACCGTGGAGGCGCAGGTGCGGTGACAGGTCCC 1800
   540  R  V  R  H  T  I  T  I  P  N  R  G  G  A  G  A  V  T  G  P     559
  1801  GACGAGTGGGCGGACTTTGGCTTTGACCTGCCTGACTGCAAGTCCCGTAAGCAGCCCATC 1860
   560  D  E  W  A  D  F  G  F  D  L  P  D  C  K  S  R  K  Q  P  I     579
  1861  AAAGAGGAGTTCACAGAGACAGAGAGCCACTGAGGAACGTACCTTCTTCTCCTGTCCTTC 1920
   580  K  E  E  F  T  E  T  E  S  H  *                               599
  1921  CTCTGTGAGAAACTGCTCTTGGAAGTGGGACCTGTTGGCTGTGCCCACAGAAACCAGCAA 1980
  1981  GGACCTTCTGCCGGATGCCATTCCTGAAGGGAAGTCGCTCATGAACTAACTCCCTCTTGG 2040
```

FIG.7

```
     1  TGGTCCCGCTTCGACCAAGACTCCGGCTACCAGCTTGCGGGCCCCGCGGAGGAGGAGACC   60
    61  CCGCTGGGGCTAGCTGGGCGACGCGCGCCAAGCGGCGGCGGGAAGGAGGCGGGAGGAGCG  120
   121  GGGCCCGAGACCCCGACTCGGGCAGAGCCAGCTGGGGAGGCGGGGCGCGCGTGGGAGCCA  180
   181  GGGGCCCGGGTGGCCGGCCCTCCTCCGCCACGGCTGAGTGCCCGCGCTGCCTTCCCGCCG  240
   241  GTCCGCCAAGAAAGGCGCTAAGCCTGCGGCAGTCCCCTCGCCGCCGCCTCCCTGCTCCGC  300
   301  ACCCTTATAACCCGCCGTCCCGCATCCAGGCGAGGAGGCAACGCTGCAGCCCAGCCCTCG  360
   361  CCGACGCCGACGCCCGGCCCGGAGCAGAATGAGCGGCAGCGTTGGGGAGATGGCCCAGAC  420
    -8                                  M  S  G  S  V  G  E  M  A  Q  T   11
   421  CTCTTCTTCCTCCTCCTCCACCTTCGAGCACCTGTGGAGTTCTCTAGAGCCAGACAGCAC  480
    12   S  S  S  S  S  S  T  F  E  H  L  W  S  S  L  E  P  D  S  T   31
   481  CTACTTTGACCTCCCCCAGCCCAGCCAAGGGACTAGCGAGGCATCAGGCAGCGAGGAGTC  540
    32   Y  F  D  L  P  Q  P  S  Q  G  T  S  E  A  S  G  S  E  E  S   51
   541  CAACATGGATGTCTTCCACCTGCAAGGCATGGCCCAGTTCAATTTGCTCAGCAGTGCCAT  600
    52   N  M  D  V  F  H  L  Q  G  M  A  Q  F  N  L  L  S  S  A  M   71
   601  GGACCAGATGGGCAGCCGTGCGGCCCCGGCGAGCCCCTACACCCCGGAGCACGCCGCCAG  660
    72   D  Q  M  G  S  R  A  A  P  A  S  P  Y  T  P  E  H  A  A  S   91
   661  CGCGCCCACCCACTCGCCCTACGCGCAGCCCAGCTCCACCTTCGACACCATGTCTCCGGC  720
    92   A  P  T  H  S  P  Y  A  Q  P  S  S  T  F  D  T  M  S  P  A  111
   721  GCCTGTCATCCCTTCCAATACCGACTACCCCGGCCCCC  758
   112   P  V  I  P  S  N  T  D  Y  P  G  P    123
```

FIG. 8

```
_ Name: sr-p70a-cos3      Len:    650  Check: 9661  Weight:  1.00
_ Name: sr-p70b-cos3      Len:    650  Check: 3605  Weight:  1.00
_ Name: sr-p70-ht29       Len:    650  Check:   85  Weight:  1.00
_ Name: sr-p70c-att20     Len:    650  Check: 4072  Weight:  1.00
_ Name: sr-p70a-att20     Len:    650  Check: 4204  Weight:  1.00
_
_//
_
_                  1                                                  50
_ sr-p70a-cos3     .......MAQ STTTSPDGGT TFEHLWSSLE PDSTYFDLPQ SSRGNNEVVG
_ sr-p70b-cos3     .......MAQ STTTSPDGGT TFEHLWSSLE PDSTYFDLPQ SSRGNNEVVG
_  sr-p70-ht29     .......MAQ STATSPDGGT TFEHLWSSLE PDSTYFDLPQ SSRGNNEVVG
_sr-p70c-att20     .......... .......... .......... .......... ..........
_sr-p70a-att20     MSGSVGEMAQ ...TSSSSSS TFEHLWSSLE PDSTYFDLPQ PSQGTSEASG
_
_                  51                                                 100
_ sr-p70a-cos3     GTDSSMD.VF HLEGMTTSVM AQFNLLSSTM DQMSSRAASA SPYTPEHAAS
_ sr-p70b-cos3     GTDSSMD.VF HLEGMTTSVM AQFNLLSSTM DQMSSRAASA SPYTPEHAAS
_  sr-p70-ht29     GTDSSMD.VF HLEGMTTSVM AQFNLLSSTM DQMSSRAASA SPYTPEHAAS
_sr-p70c-att20     ...MCMGPVY ..ESLG...Q AQFNLLSSAM DQMGSRAAPA SPYTPEHAAS
_sr-p70a-att20     SEESNMD.VF HLQGM..... AQFNLLSSAM DQMGSRAAPA SPYTPEHAAS
_
_                  101                                                150
_ sr-p70a-cos3     VPTHSPYAQP SSTFDTMSPA PVIPSNTDYP GPHHFEVTFQ QSSTAKSATW
_ sr-p70b-cos3     VPTHSPYAQP SSTFDTMSPA PVIPSNTDYP GPHHFEVTFQ QSSTAKSATW
_  sr-p70-ht29     VPTHSPYAQP SSTFDTMSPA PVIPSNTDYP GPHHFEVTFQ QSSTAKSATW
_sr-p70c-att20     APTHSPYAQP SSTFDTMSPA PVIPSNTDYP GPHHFEVTFQ QSSTAKSATW
_sr-p70a-att20     APTHSPYAQP SSTFDTMSPA PVIPSNTDYP GP........ ..........
_
_                  151                                                200
_ sr-p70a-cos3     TYSPLLKKLY CQIAKTCPIQ IKVSAPPPPG TAIRAMPVYK KAEHVTDIVK
_ sr-p70b-cos3     TYSPLLKKLY CQIAKTCPIQ IKVSAPPPPG TAIRAMPVYK KAEHVTDIVK
_  sr-p70-ht29     TYSPLLKKLY CQIAKTCPIQ IKVSTPPPPG TAIRAMPVYK KAEHVTDIVK
_sr-p70c-att20     TYSPLLKKLY CQIAKTCPIQ IKVSTPPPPG TAIRAMPVYK KAEHVTDIVK
_sr-p70a-att20     .......... .......... .......... .......... ..........
_
_                  201                                                250
_ sr-p70a-cos3     RCPNHELGRD FNEGQSAPAS HLIRVEGNNL SQYVDDPVTG RQSVVVPYEP
_ sr-p70b-cos3     RCPNHELGRD FNEGQSAPAS HLIRVEGNNL SQYVDDPVTG RQSVVVPYEP
_  sr-p70-ht29     RCPNHELGRD FNEGQSAPAS HLIRVEGNNL SQYVDDPVTG RQSVVVPYEP
_sr-p70c-att20     RCPNHELGRD FNEGQSAPAS HLIRVEGNNL AQYVDDPVTG RQSVVVPYEP
_sr-p70a-att20     .......... .......... .......... .......... ..........
_
_                  251                                                300
_ sr-p70a-cos3     PQVGTEFTTI LYNFMCNSSC VGGMNRRPIL IIITLETRDG QVLGRRSFEG
_ sr-p70b-cos3     PQVGTEFTTI LYNFMCNSSC VGGMNRRPIL IIITLETRDG QVLGRRSFEG
_  sr-p70-ht29     PQVGTEFTTI LYNFMCNSSC VGGMNRRPIL IIITLEMRDG QVLGRRSFEG
_sr-p70c-att20     PQVGTEFTTI LYNFMCNSSC VGGMNRRPIL VIITLETRDG QVLGRRSFEG
_sr-p70a-att20     .......... .......... .......... .......... ..........
_
_                  301                                                350
_ sr-p70a-cos3     RICACPGRDR KADEDHYREQ QALNESSAKN GAASKRAFKQ SPPAVPALGP
_ sr-p70b-cos3     RICACPGRDR KADEDHYREQ QALNESSAKN GAASKRAFKQ SPPAVPALGP
_  sr-p70-ht29     RICACPGRDR KADEDHYREQ QALNESSAKN GAASKRAFKQ SPPAVPALGA
_sr-p70c-att20     RICACPGRDR KADEDHYREQ QALNESTTKN GAASKRAFKQ SPPAIPALGT
_sr-p70a-att20     .......... .......... .......... .......... ..........
_
```

FIG.9

FIG. 9A

```
                351                                                     400
 sr-p70a-cos3   GVKKRRHGDE DTYYLQVRGR ENFEILMKLK ESLELMELVP QPLVDSYR..
 sr-p70b-cos3   GVKKRRHGDE DTYYLQVRGR ENFEILMKLK ESLELMELVP QPLVDSYR..
   sr-p70-ht29  GVKKRRHGDE DTYYLQVRGR ENFEILMKLK ESLELMELVP QPLVDSYR..
_sr-p70c-att20  NVKKRRHGDE DMFYMHVRGR ENFEILMKVK ESLELMELVP QPLVDSYRQQ
_sr-p70a-att20  .......... .......... .......... .......... ..........

                401                                                     450
 sr-p70a-cos3   QQQQLLQRPS HLQPPSYGPV LSPMNKVHGG VNKLPSVNQL VGQPPPHSSA
 sr-p70b-cos3   QQQQLLQRPS HLQPPSYGPV LSPMNKVHGG VNKLPSVNQL VGQPPPHSSA
   sr-p70-ht29  QQQQLLQRPS HLQPPSYGPV LSPMNKVHGG MNKLPSVNQL VGQPPPHSSA
_sr-p70c-att20  QQQQLLQRPS HLQPPSYGPV LSPMNKVHGG VNKLPSVNQL VGQPPPHSSA
_sr-p70a-att20  .......... .......... .......... .......... ..........

                451                                                     500
 sr-p70a-cos3   ATPNLGPVGS GMLNNHGHAV PANSEMTSSH GTQSMVSGSH CTPPPPYHAD
 sr-p70b-cos3   ATPNLGPVGS GMLNNHGHAV PANSEMTSSH GTQSMVSGSH CTPPPPYHAD
   sr-p70-ht29  ATPNLGPVGP GMLNNHGHAV PANGEMSSSH SAQSMVSGSH CTPPPPYHAD
_sr-p70c-att20  AGPNLGPMGS GMLNSHGHSM PANGEMNGGH SSQTMVSGSH CTPPPPYHAD
_sr-p70a-att20  .......... .......... .......... .......... ..........

                501                                                     550
 sr-p70a-cos3   PSLVSFLTGL GCPNCIEYFT SQGLQSIYHL QNLTIEDLGA LKIPEQYRMT
 sr-p70b-cos3   PSLVR..T.W G.P..... .......... .......... ..........
   sr-p70-ht29  PSLVSFLTGL GCPNCIEYFT SQGLQSIYHL QNLTIEDLGA LKIPEQYRMT
_sr-p70c-att20  PSLVSFLTGL GCPNCIECFT SQGLQSIYHL QNLTIEDLGA LKVPDQYRMT
_sr-p70a-att20  .......... .......... .......... .......... ..........

                551                                                     600
 sr-p70a-cos3   IWRGLQDLKQ GHDYGAAAQQ LLR.SSNAAA ISIGGSGELQ RQRVMEAVHF
 sr-p70b-cos3   .......... .......... .......... .......... ..........
   sr-p70-ht29  IWRGLQDLKQ GHDYS.TAQQ LLR.SSNAAT ISIGGSGELQ RQRVMEAVHF
_sr-p70c-att20  IWRGLQDLKQ SHDCG...QQ LLRSSSNAAT ISIGGSGELQ RQRVMEAVHF
_sr-p70a-att20  .......... .......... .......... .......... ..........

                601                                                     650
 sr-p70a-cos3   RVRHTITIPN RGGPGA..GP DEWADFGFDL PDCKARKQPI KEEFTEAEIH
 sr-p70b-cos3   .......... .......... .......... .......... ..........
   sr-p70-ht29  RVRHTITIPN RGGPGG..GP DEWADFGFDL PDCKARKQPI KEEFTEAEIH
_sr-p70c-att20  RVRHTITIPN RGGAGAVTGP DEWADFGFDL PDCKSRKQPI KEEFTETESH
_sr-p70a-att20  .......... .......... .......... .......... ..........
```

FIG. 9B

~~FIG. 9 cont.~~

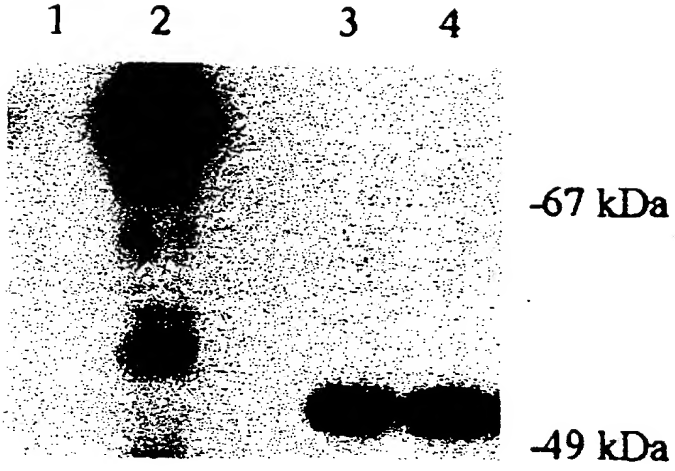

FIG.10a

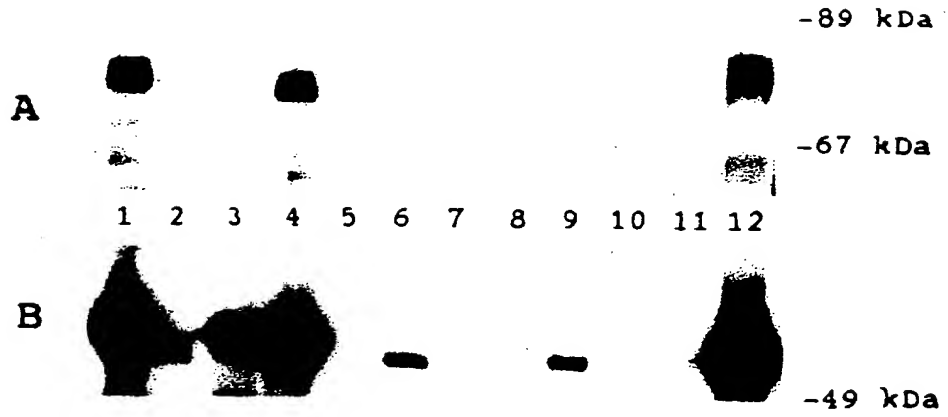

FIG.10b

FIG.11

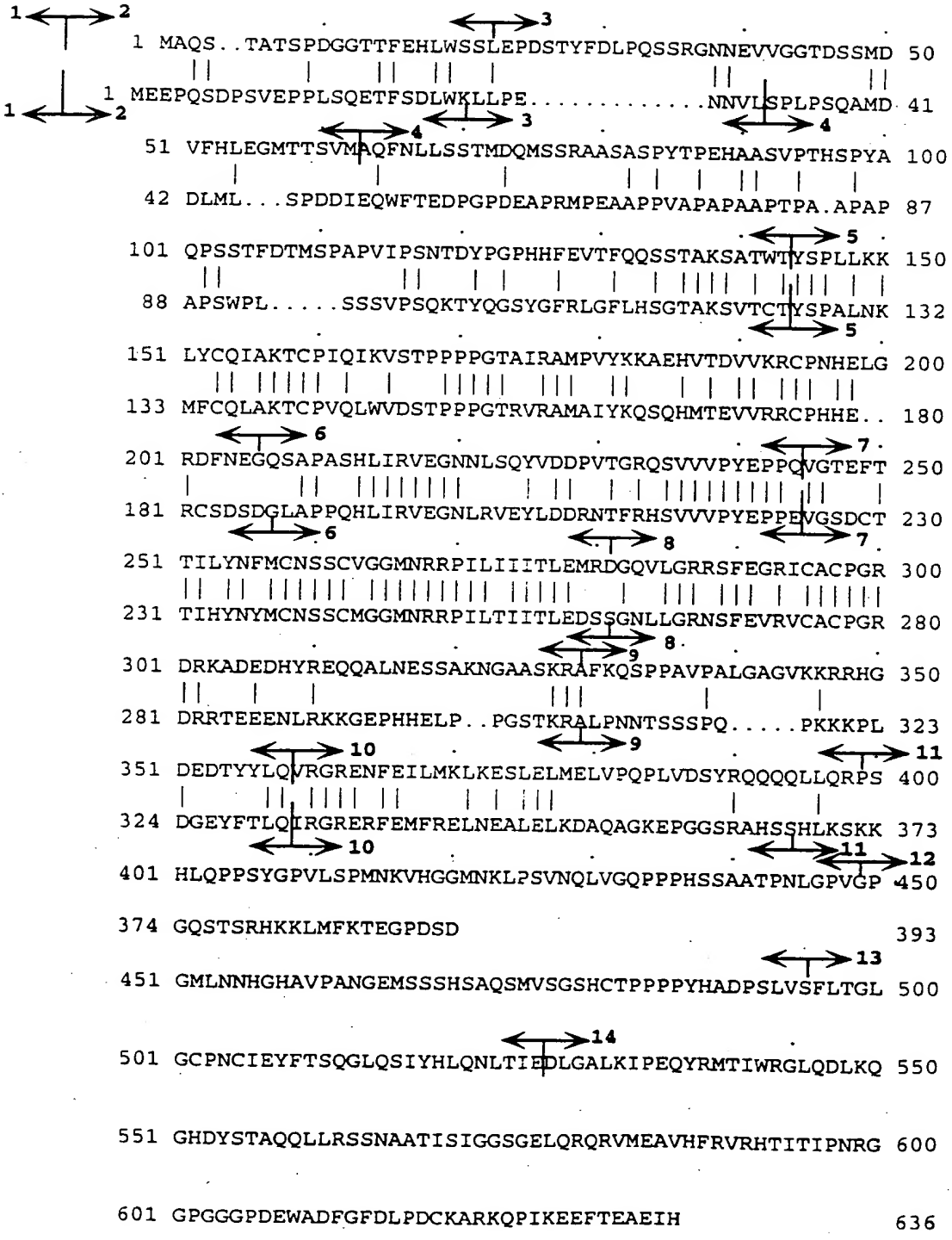
1 MAQS..TATSPDGGTTFEHLWSSLEPDSTYFDLPQSSRGNNEVVGGTDSSMD 50
1 MEEPQSDPSVEPPLSQETFSDLWKLLPE.....NNVLSPLPSQAMD 41
51 VFHLEGMTTSVMAQFNLLSSTMDQMSSRAASASPYTPEHAASVPTHSPYA 100
42 DLML...SPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPA.APAP 87
101 QPSSTFDTMSPAPVIPSNTDYPGPHHFEVTFQQSSTAKSATWTYSPLLKK 150
88 APSWPL.....SSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNK 132
151 LYCQIAKTCPIQIKVSTPPPPGTAIRAMPVYKKAEHVTDVVKRCPNHELG 200
133 MFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE.. 180
201 RDFNEGQSAPASHLIRVEGNNLSQYVDDPVTGRQSVVVPYEPPQVGTEFT 250
181 RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCT 230
251 TILYNFMCNSSCVGGMNRRPILIIITLEMRDGQVLGRRSFEGRICACPGR 300
231 TIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGR 280
301 DRKADEDHYREQQALNESSAKNGAASKRAFKQSPPAVPALGAGVKKRRHG 350
281 DRRTEEENLRKKGEPHHELP..PGSTKRALPNNTSSSPQ.....PKKKPL 323
351 DEDTYYLQVRGRENFEILMKLKESLELMELVPQPLVDSYRQQQQLLQRPS 400
324 DGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKK 373
401 HLQPPSYGPVLSPMNKVHGGMNKLPSVNQLVGQPPPHSSAATPNLGPVGP 450
374 GQSTSRHKKLMFKTEGPDSD 393
451 GMLNNHGHAVPANGEMSSSHSAQSMVSGSHCTPPPPYHADPSLVSFLTGL 500
501 GCPNCIEYFTSQGLQSIYHLQNLTIEDLGALKIPEQYRMTIWRGLQDLKQ 550
551 GHDYSTAQQLLRSSNAATISIGGSGELQRQRVMEAVHFRVRHTITIPNRG 600
601 GPGGGPDEWADFGFDLPDCKARKQPIKEEFTEAEIH 636

FIG. 12

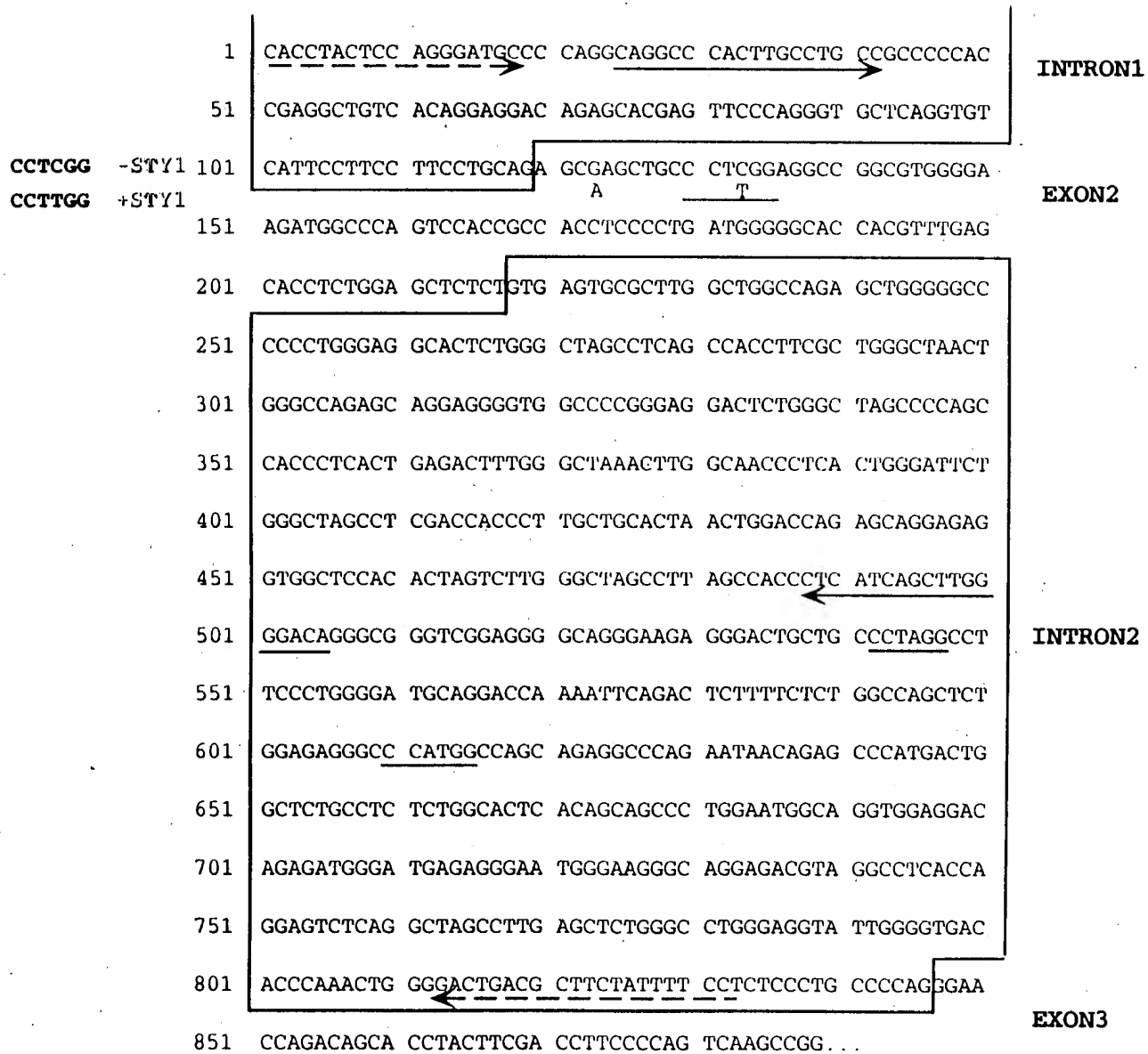
1 CACCTACTCC AGGGATGCCC CAGGCAGGCC CACTTGCCTG CCGCCCCCAC
INTRON1
51 CGAGGCTGTC ACAGGAGGAC AGAGCACGAG TTCCCAGGGT GCTCAGGTGT
CCTCGG -STY1 101 CATTCCTTCC TTCCTGCAGA GCGAGCTGCC CTCGGAGGCC GGCGTGGGGA
CCTTGG +STY1
A T
EXON2
151 AGATGGCCCA GTCCACCGCC ACCTCCCCTG ATGGGGGCAC CACGTTTGAG
201 CACCTCTGGA GCTCTCTGTG AGTGCGCTTG GCTGGCCAGA GCTGGGGGCC
251 CCCCTGGGAG GCACTCTGGG CTAGCCTCAG CCACCTTCGC TGGGCTAACT
301 GGGCCAGAGC AGGAGGGGTG GCCCCGGGAG GACTCTGGGC TAGCCCCAGC
351 CACCCTCACT GAGACTTTGG GCTAAACTTG GCAACCCTCA CTGGGATTCT
401 GGGCTAGCCT CGACCACCCT TGCTGCACTA ACTGGACCAG AGCAGGAGAG
451 GTGGCTCCAC ACTAGTCTTG GGCTAGCCTT AGCCACCCTC ATCAGCTTGG
501 GGACAGGGCG GGTCGGAGGG GCAGGGAAGA GGGACTGCTG CCCTAGGCCT
INTRON2
551 TCCCTGGGGA TGCAGGACCA AAATTCAGAC TCTTTTCTCT GGCCAGCTCT
601 GGAGAGGGCC CATGGCCAGC AGAGGCCCAG AATAACAGAG CCCATGACTG
651 GCTCTGCCTC TCTGGCACTC ACAGCAGCCC TGGAATGGCA GGTGGAGGAC
701 AGAGATGGGA TGAGAGGGAA TGGGAAGGGC AGGAGACGTA GGCCTCACCA
751 GGAGTCTCAG GCTAGCCTTG AGCTCTGGGC CTGGGAGGTA TTGGGGTGAC
801 ACCCAAACTG GGGACTGACG CTTCTATTTT CCTCTCCCTG CCCCAGGGAA
EXON3
851 CCAGACAGCA CCTACTTCGA CCTTCCCCAG TCAAGCCGG...

FIG.13

```
sr-p70d-imr32                        CG ACCTTCCCCA GTCAAGCCGG GGGAATAATG 32
 sr-p70a-ht29                        CG ACCTTCCCCA GTCAAGCCGG GGGAATAATG 150

AGGTGGTGGG CGGAACGGAT TCCAGCATGG ACGTCTTCCA CCTGGAGGGC 82
AGGTGGTGGG CGGAACGGAT TCCAGCATGG ACGTCTTCCA CCTGGAGGGC 200

ATGACTACAT CTGTCATGCA TCCTCGGCTC CTGCCTCACT AGCTGCGGAG 132
ATGACTACAT CTGTCAT... .......... .......... .......... 217

CCTCTCCCGC TCGGTCCACG CTGCCGGGCG GCCACGACCG TGACCCTTCC 182
.......... .......... .......... .......... ..........

CCTCGGGCCG CCCAGATCCA TGCCTCGTCC CACGGGACAC CAGTTCCCTG 232
.......... .......... .......... .......... ..........

GCGTGTGCAG ACCCCCCGGC GCCTACCATG CTGTACGTCG GTGACCCCGC 282
.......... .......... .......... .......... ..........

ACGGCACCTC GCCACGGCCC AGTTCAATCT GCTGAGCAGC ACCATGGACC 332
.......... .....GGCCC AGTTCAATCT GCTGAGCAGC ACCATGGACC 252

AGATGAGCAG CCGCGCGGCC TCGGCCAGCC CCTACACCCC AGAGCACGCC 382
AGATGAGCAG CCGCGCGGCC TCGGCCAGCC CCTACACCCC AGAGCACGCC 302

GCCAGCGTGC CCACCCACTC GCCCTACGCA CAACCCAGCT CCACCTTCGA 432
GCCAGCGTGC CCACCCAcTC GCCCTACGCA CAACCCAGCT CCACCTTCGA 352

CACCATGTCG CCGGCGCCTG TCATCCCCTC CAACACCGAC TACCCCGGAC 482
CACCATGTCG CCGGCGCCTG TCATCCCCTC CAACACCGAC TACCCCGGAC 402

CCCACCACTT TGAGGTCACT TTCCAGCAGT CCAGCACGGC CAAGTCAGCC 532
CCCACCACTT TGAGGTCACT TTCCAGCAGT CCAGCACGGC CAAGTCAGCC 452

ACCTGGACGT ACTCCCCGCT CTTGAAG
ACCTGGACGT ACTCCCCGCT CTTGAAG
```

FIG. 14

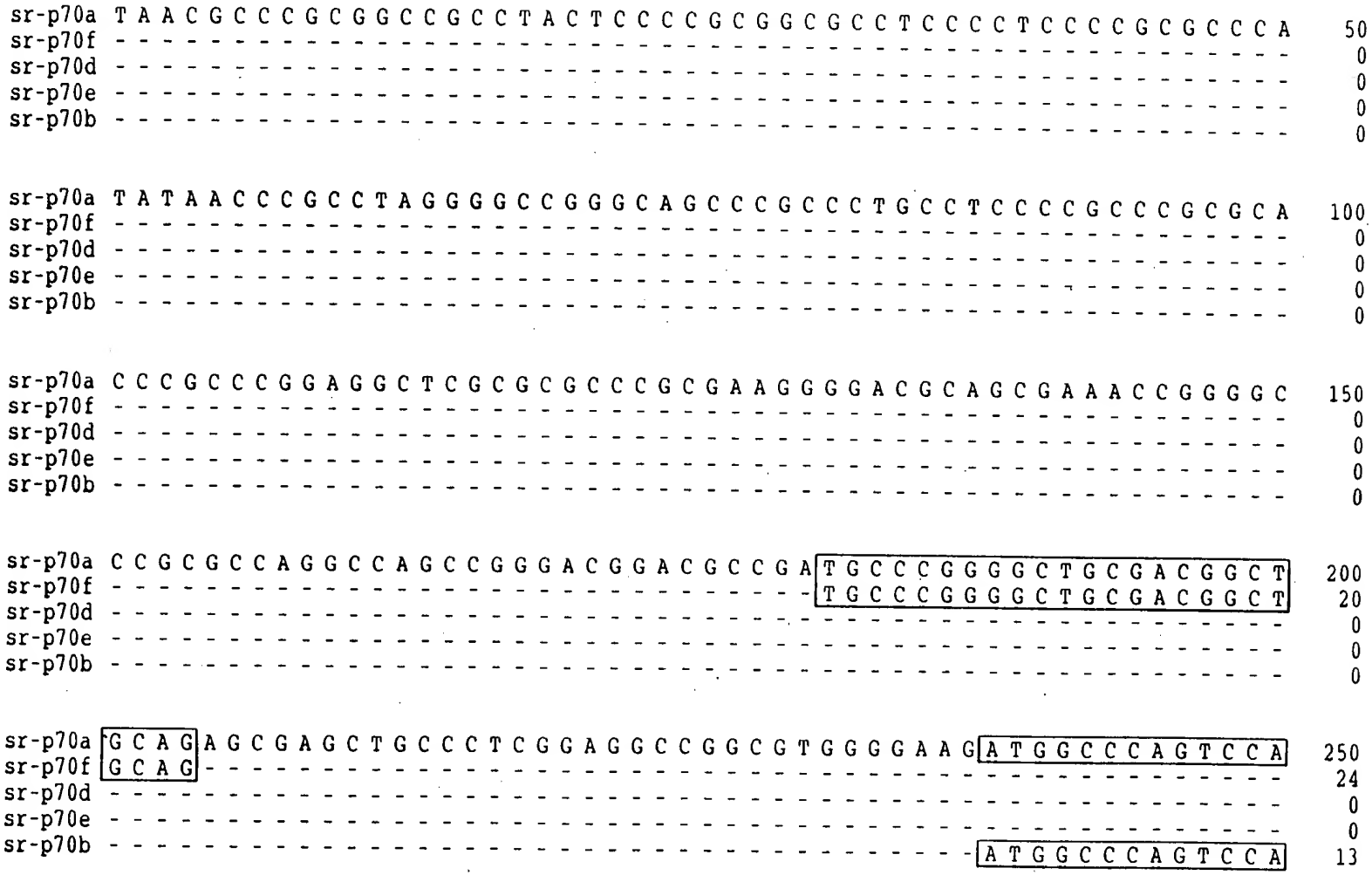

FIG. 15A

~~FIG. 15~~

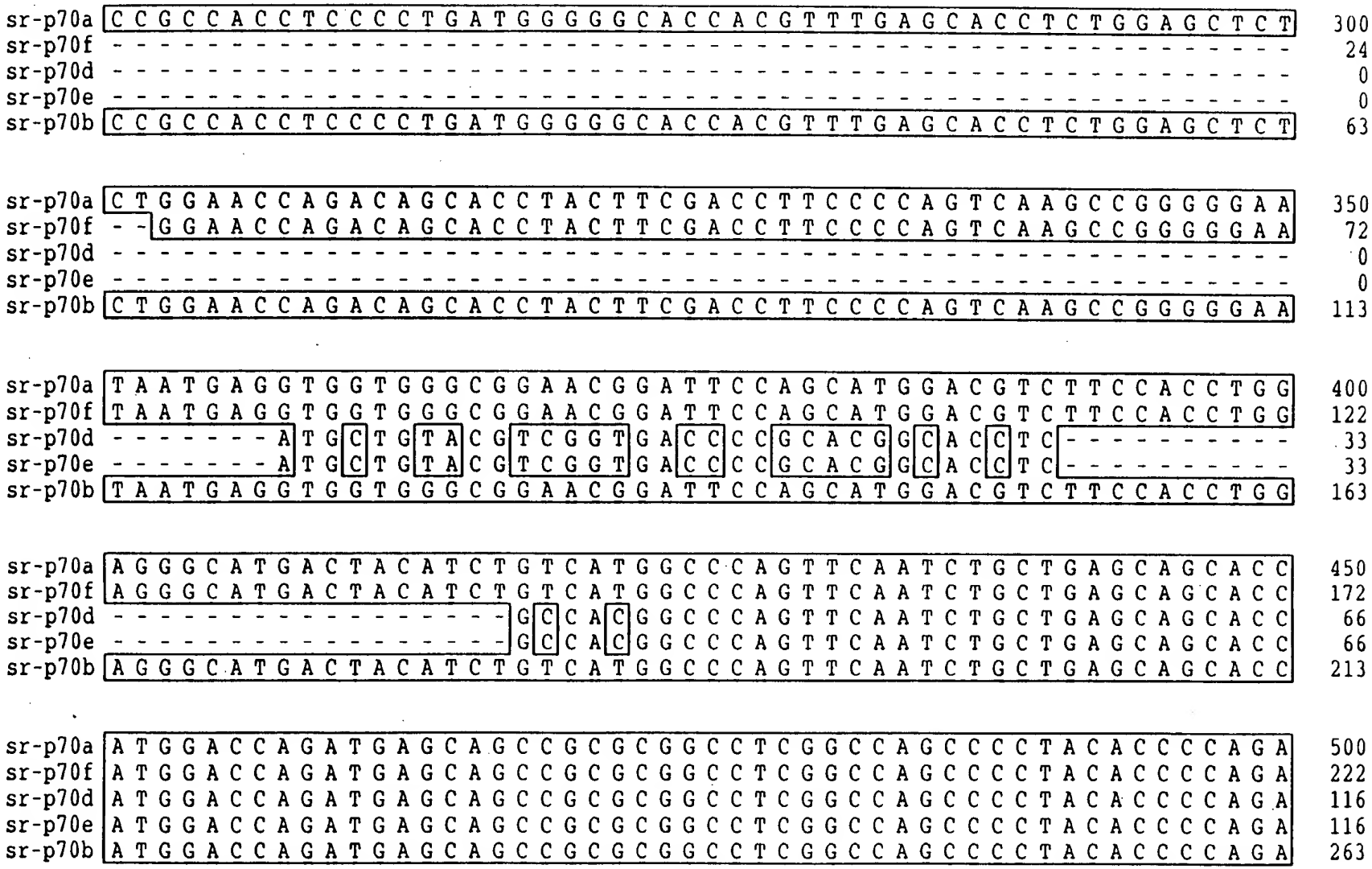

FIG. 15B

~~FIG. 15 cont.~~

```
sr-p70a GCACGCCGCCAGCGTGCCCACCCACTCGCCCTACGCACAACCCAGCTCCA 550
sr-p70f GCACGCCGCCAGCGTGCCCACCCACTCGCCCTACGCACAACCCAGCTCCA 272
sr-p70d GCACGCCGCCAGCGTGCCCACCCACTCGCCCTACGCACAACCCAGCTCCA 166
sr-p70e GCACGCCGCCAGCGTGCCCACCCACTCGCCCTACGCACAACCCAGCTCCA 166
sr-p70b GCACGCCGCCAGCGTGCCCACCCACTCGCCCTACGCACAACCCAGCTCCA 313

sr-p70a CCTTCGACACCATGTCGCCGGCGCCTGTCATCCCCTCCAACACCGACTAC 600
sr-p70f CCTTCGACACCATGTCGCCGGCGCCTGTCATCCCCTCCAACACCGACTAC 322
sr-p70d CCTTCGACACCATGTCGCCGGCGCCTGTCATCCCCTCCAACACCGACTAC 216
sr-p70e CCTTCGACACCATGTCGCCGGCGCCTGTCATCCCCTCCAACACCGACTAC 216
sr-p70b CCTTCGACACCATGTCGCCGGCGCCTGTCATCCCCTCCAACACCGACTAC 363

sr-p70a CCCGGACCCCACCACTTTGAGGTCACTTTCCAGCAGTCCAGCACGGCCAA 650
sr-p70f CCCGGACCCCACCACTTTGAGGTCACTTTCCAGCAGTCCAGCACGGCCAA 372
sr-p70d CCCGGACCCCACCACTTTGAGGTCACTTTCCAGCAGTCCAGCACGGCCAA 266
sr-p70e CCCGGACCCCACCACTTTGAGGTCACTTTCCAGCAGTCCAGCACGGCCAA 266
sr-p70b CCCGGACCCCACCACTTTGAGGTCACTTTCCAGCAGTCCAGCACGGCCAA 413

sr-p70a GTCAGCCACCTGGACGTACTCCCCGCTCTTGAAGAAACTCTACTGCCAGA 700
sr-p70f GTCAGCCACCTGGACGTACTCCCCGCTCTTGAAGAAACTCTACTGCCAGA 422
sr-p70d GTCAGCCACCTGGACGTACTCCCCGCTCTTGAAGAAACTCTACTGCCAGA 316
sr-p70e GTCAGCCACCTGGACGTACTCCCCGCTCTTGAAGAAACTCTACTGCCAGA 316
sr-p70b GTCAGCCACCTGGACGTACTCCCCGCTCTTGAAGAAACTCTACTGCCAGA 463

sr-p70a TCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCACCCCGCCACCCCCA 750
sr-p70f TCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCACCCCGCCACCCCCA 472
sr-p70d TCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCACCCCGCCACCCCCA 366
sr-p70e TCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCACCCCGCCACCCCCA 366
sr-p70b TCGCCAAGACATGCCCCATCCAGATCAAGGTGTCCACCCCGCCACCCCCA 513
```

FIG. 15C

~~FIG. 15 cont.~~

```
sr-p70a GGCACTGCCATCCGGGCCATGCCTGTTTACAAGAAAGCGGAGCACGTGAC  800
sr-p70f GGCACTGCCATCCGGGCCATGCCTGTTTACAAGAAAGCGGAGCACGTGAC  522
sr-p70d GGCACTGCCATCCGGGCCATGCCTGTTTACAAGAAAGCGGAGCACGTGAC  416
sr-p70e GGCACTGCCATCCGGGCCATGCCTGTTTACAAGAAAGCGGAGCACGTGAC  416
sr-p70b GGCACTGCCATCCGGGCCATGCCTGTTTACAAGAAAGCGGAGCACGTGAC  563

sr-p70a CGACGTCGTGAAACGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACG  850
sr-p70f CGACGTCGTGAAACGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACG  572
sr-p70d CGACGTCGTGAAACGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACG  466
sr-p70e CGACGTCGTGAAACGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACG  466
sr-p70b CGACGTCGTGAAACGCTGCCCCAACCACGAGCTCGGGAGGGACTTCAACG  613

sr-p70a AAGGACAGTCTGCTCCAGCCAGCCACCTCATCCGCGTGGAAGGCAATAAT  900
sr-p70f AAGGACAGTCTGCTCCAGCCAGCCACCTCATCCGCGTGGAAGGCAATAAT  622
sr-p70d AAGGACAGTCTGCTCCAGCCAGCCACCTCATCCGCGTGGAAGGCAATAAT  516
sr-p70e AAGGACAGTCTGCTCCAGCCAGCCACCTCATCCGCGTGGAAGGCAATAAT  516
sr-p70b AAGGACAGTCTGCTCCAGCCAGCCACCTCATCCGCGTGGAAGGCAATAAT  663

sr-p70a CTCTCGCAGTATGTGGATGACCCTGTCACCGGCAGGCAGAGCGTCGTGGT  950
sr-p70f CTCTCGCAGTATGTGGATGACCCTGTCACCGGCAGGCAGAGCGTCGTGGT  672
sr-p70d CTCTCGCAGTATGTGGATGACCCTGTCACCGGCAGGCAGAGCGTCGTGGT  566
sr-p70e CTCTCGCAGTATGTGGATGACCCTGTCACCGGCAGGCAGAGCGTCGTGGT  566
sr-p70b CTCTCGCAGTATGTGGATGACCCTGTCACCGGCAGGCAGAGCGTCGTGGT  713

sr-p70a GCCCTATGAGCCACCACAGGTGGGGACGGAATTCACCACCATCCTGTACA 1000
sr-p70f GCCCTATGAGCCACCACAGGTGGGGACGGAATTCACCACCATCCTGTACA  722
sr-p70d GCCCTATGAGCCACCACAGGTGGGGACGGAATTCACCACCATCCTGTACA  616
sr-p70e GCCCTATGAGCCACCACAGGTGGGGACGGAATTCACCACCATCCTGTACA  616
sr-p70b GCCCTATGAGCCACCACAGGTGGGGACGGAATTCACCACCATCCTGTACA  763
```

FIG. 15D

~~FIG. 15 cont.~~

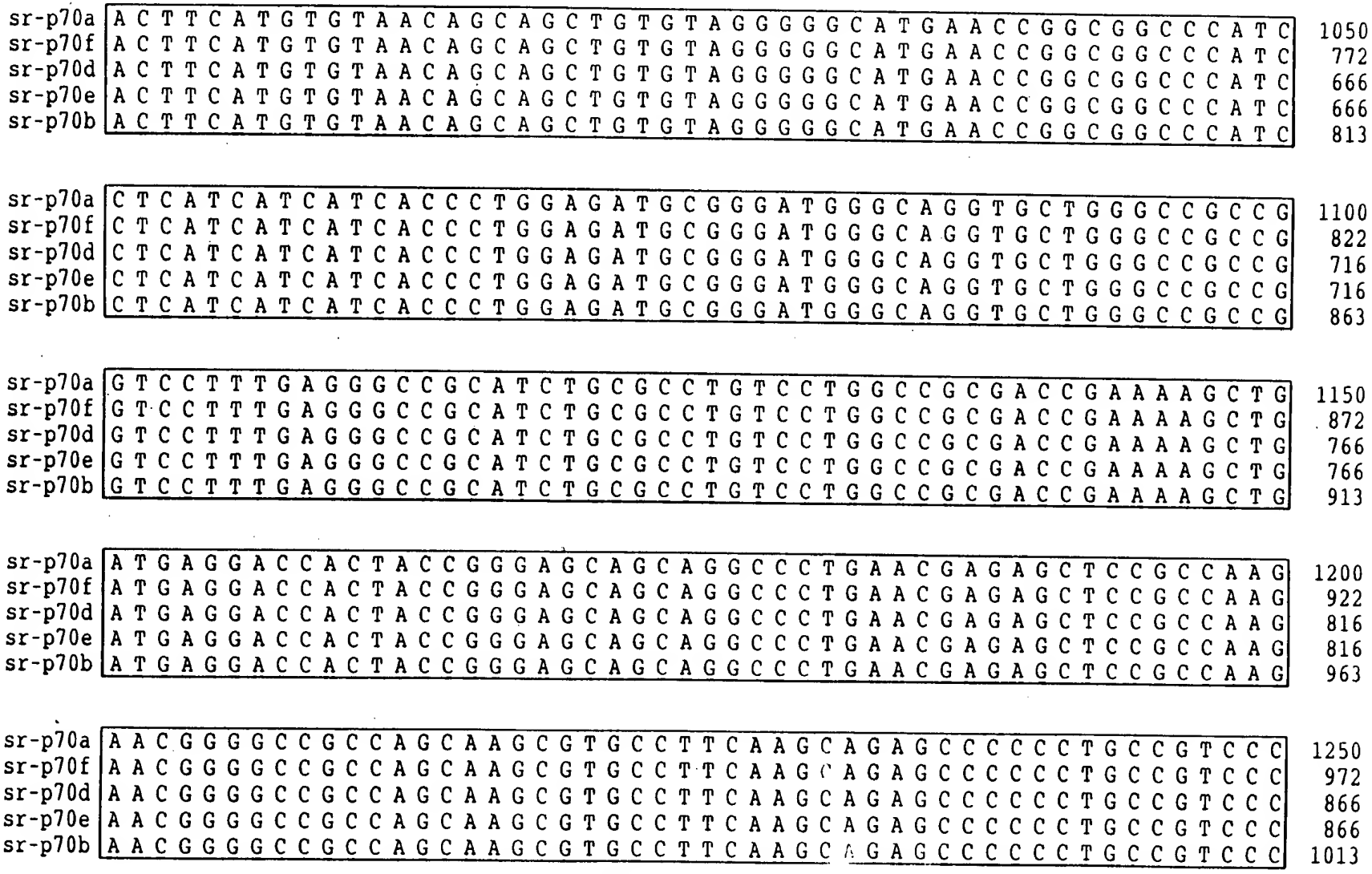

FIG. 15E

~~FIG. 15 cont.~~

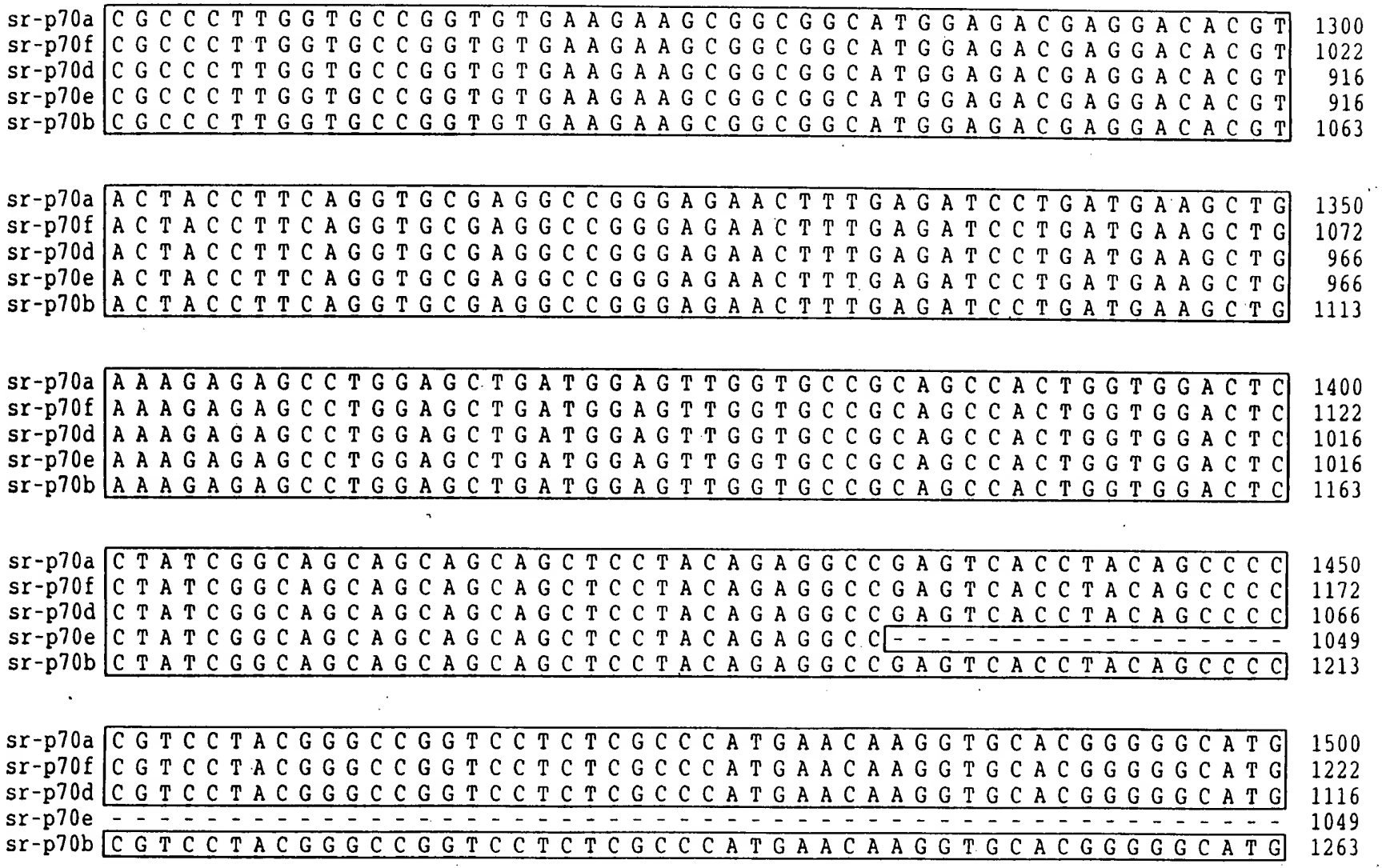

FIG. 15F

~~FIG. 15 cont.~~

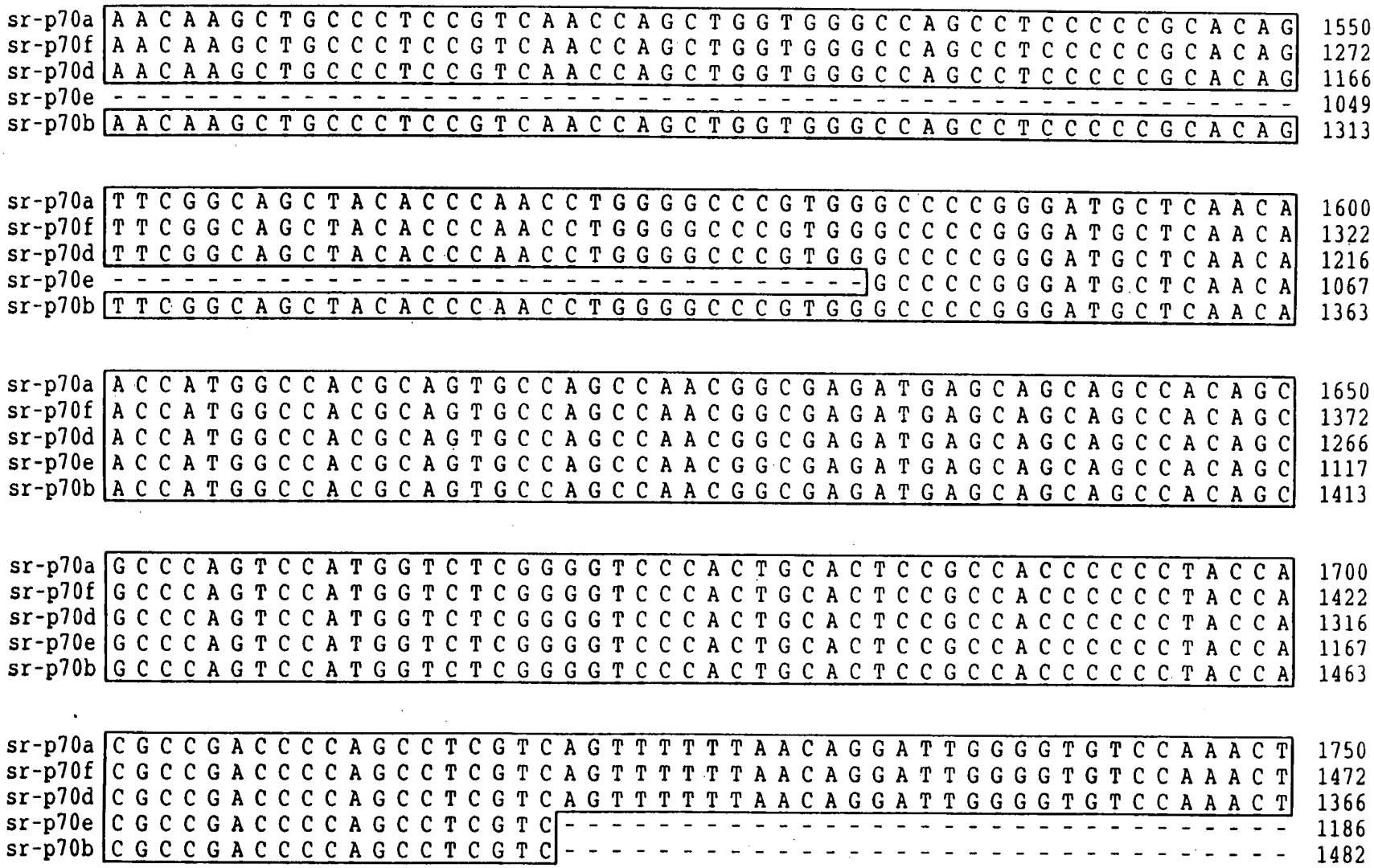

FIG. 15G

FIG. 15 cont.

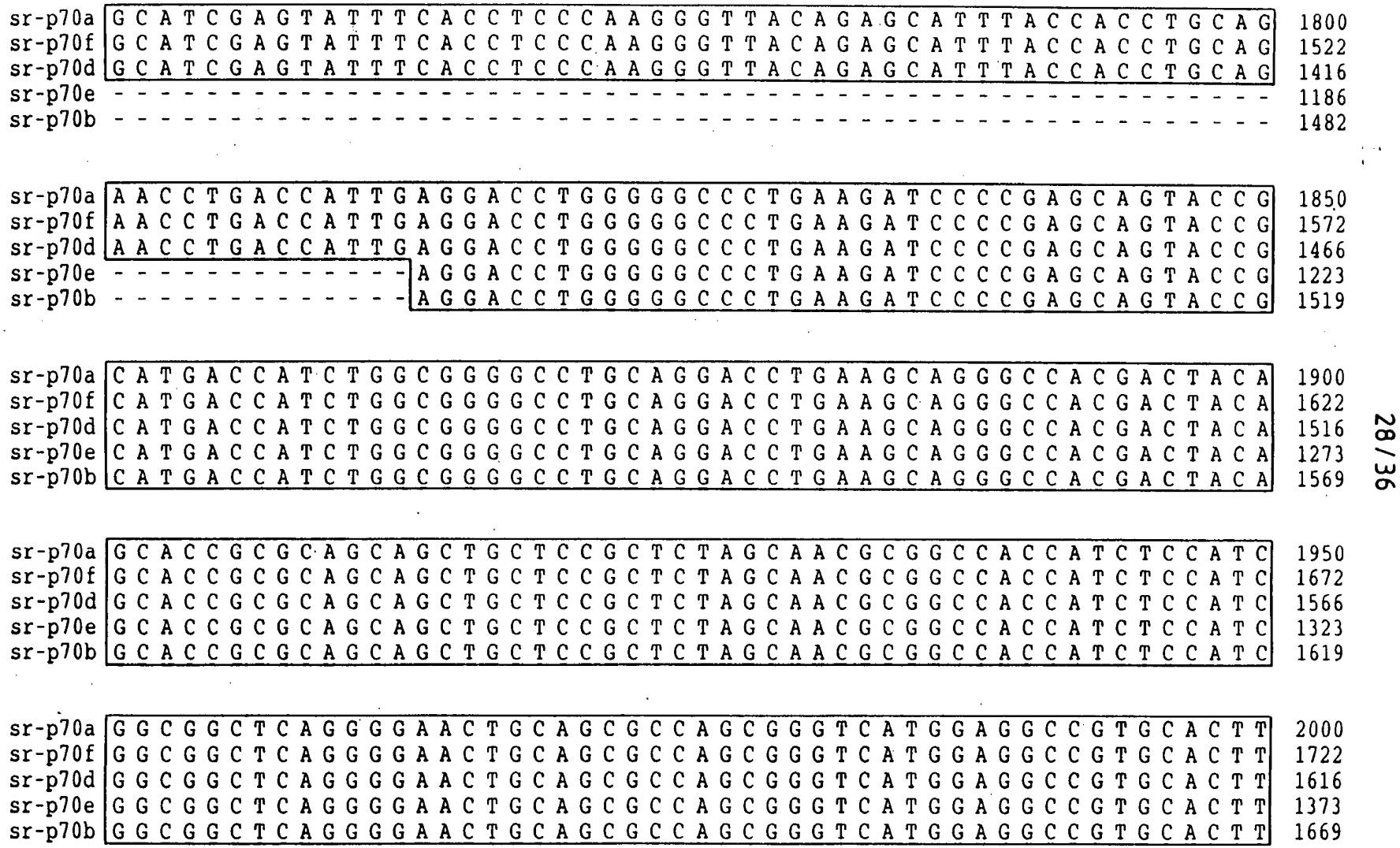

FIG. 15H

~~FIG. 15cont.~~

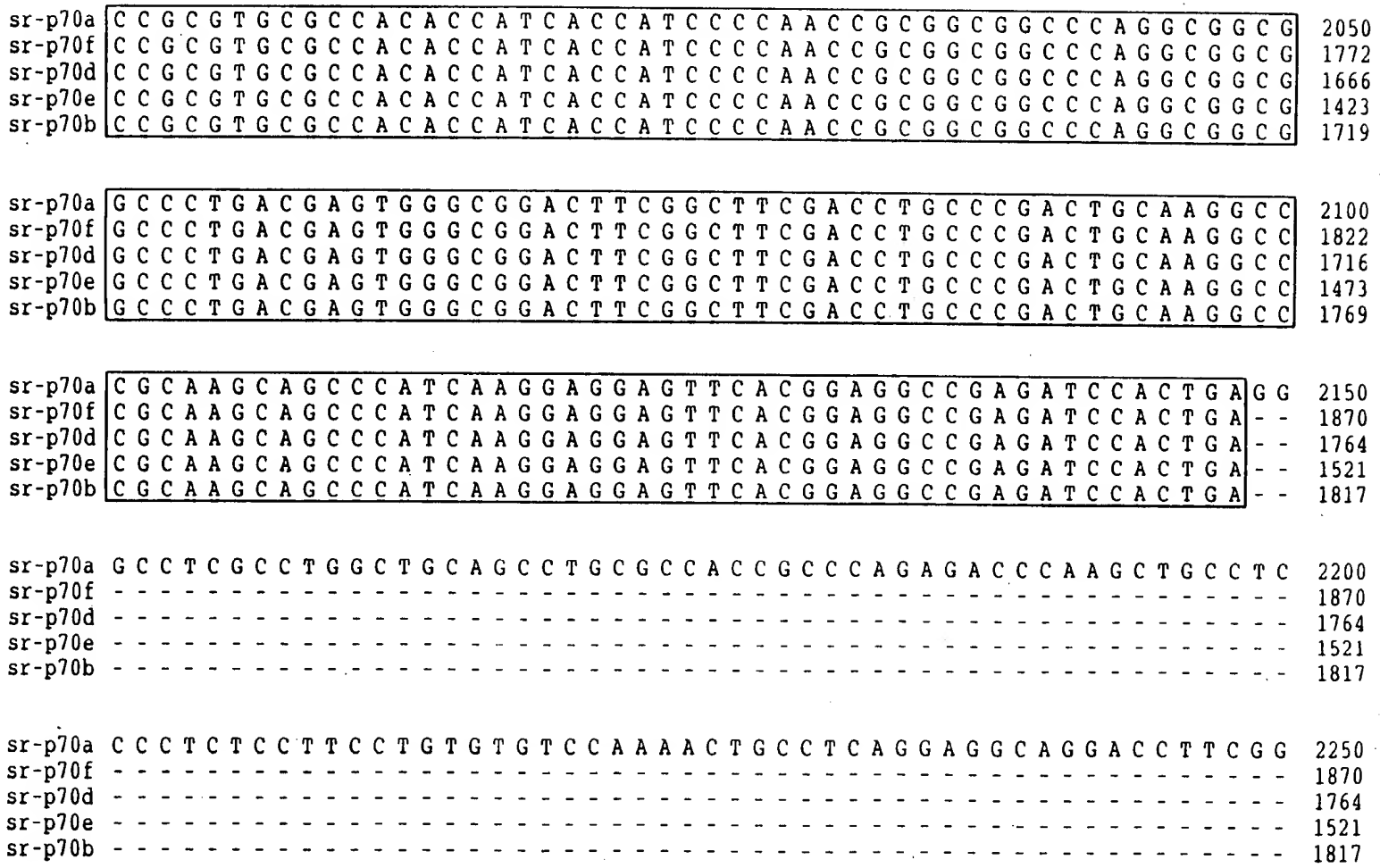

FIG. 15I

~~FIG. 15 cont.~~

```
sr-p70a G C T G T G C C C G G G G A A A G G C A A G G T C C G G C C C A T C C C C A G G C A C C T C A C A G  2300
sr-p70f - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1870
sr-p70d - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1764
sr-p70e - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1521
sr-p70b - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1817

sr-p70a G C C C C A G G A A A G G C C C A G C C A C C G A A G C C G C C T G T G G A C A G C C T G A G T C A  2350
sr-p70f - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1870
sr-p70d - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1764
sr-p70e - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1521
sr-p70b - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1817

sr-p70a C C T G C A G A A C C                                                                           2361
sr-p70f - - - - - - - - - - -                                                                           1870
sr-p70d - - - - - - - - - - -                                                                           1764
sr-p70e - - - - - - - - - - -                                                                           1521
sr-p70b - - - - - - - - - - -                                                                           1817
```


FIG. 15J

~~FIG. 15 cont.~~

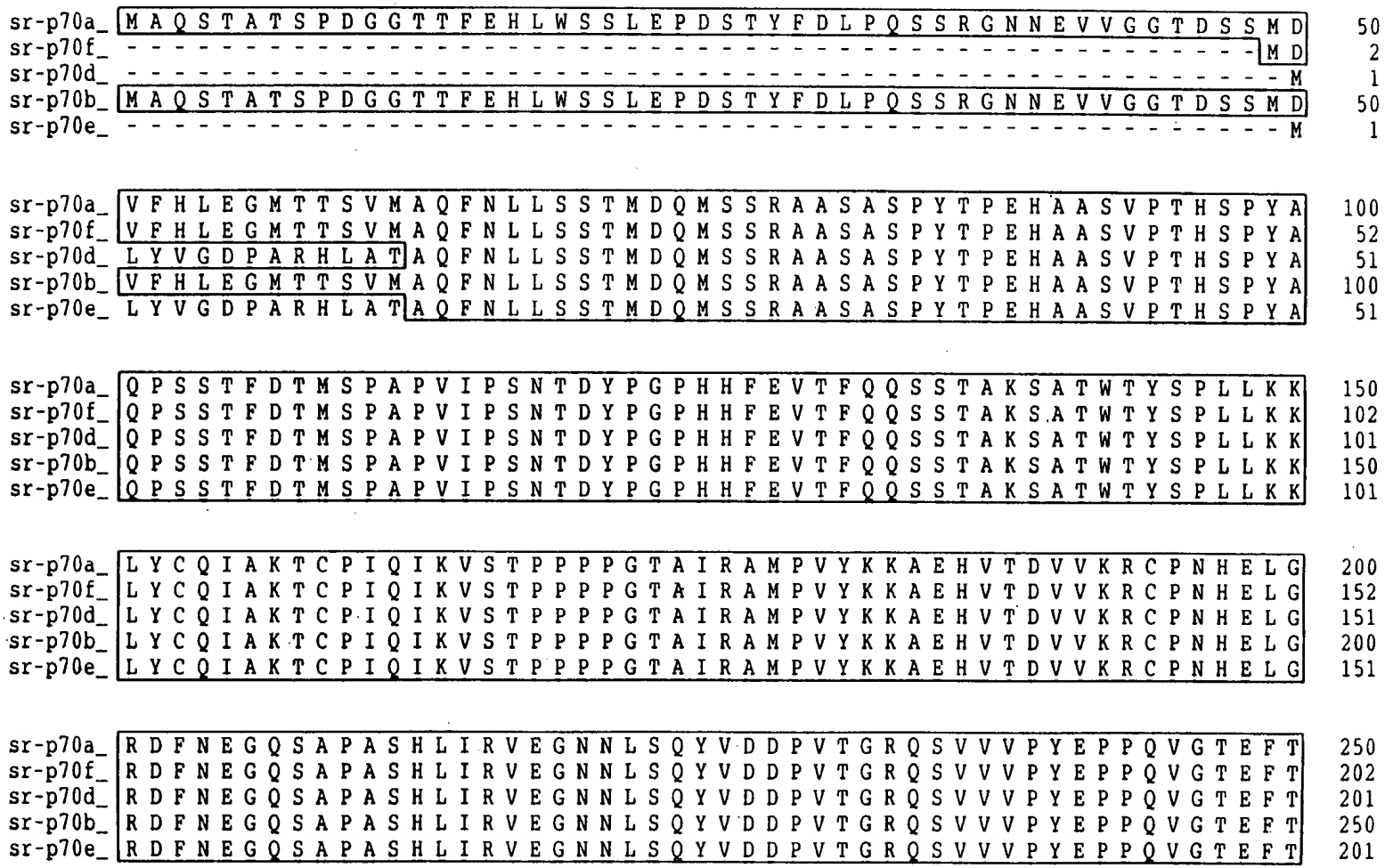

FIG. 16A

~~FIG. 16~~

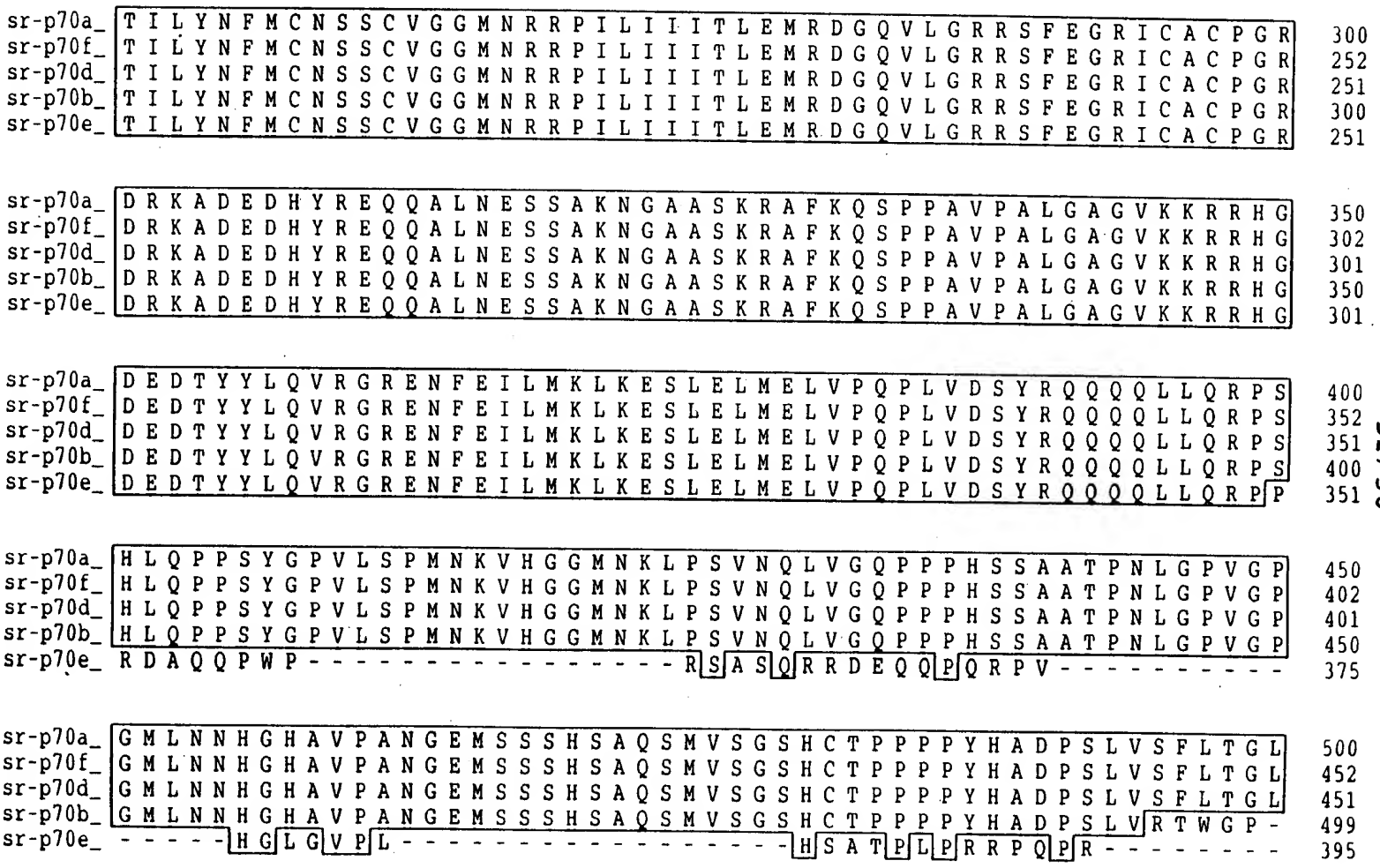

FIG. 16B

~~FIG. 16 cont.~~

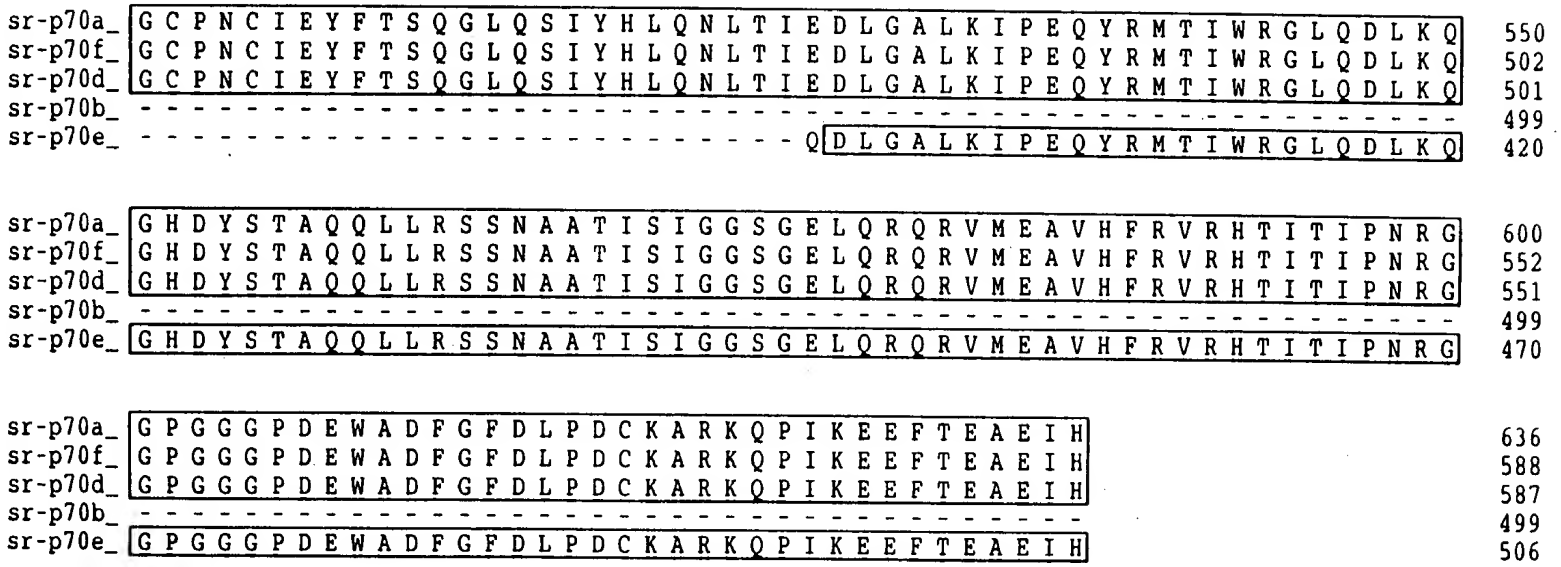

FIG. 16a

FIG.16 cont.

```
  1  TAACGCCCGCGGCCGCCTACTCCCCGCGGCGCCTCCCCTCCCCGCGCCCATATAACCCGC   60
 61  CTAGGGGCCGGGCAGCCCGCCCTGCCTCCCCGCCCGCGCACCCGCCCGGAGGCTCGCGCG  120
121  CCCGCGAAGGGGACGCAGCGAAACCGGGGCCCGCGCCAGGCCAGCCGGGACGGACGCCGA  180
181  TGCCCGGGGCTGCGACGGCTGCAGAGCAAGCTGCCCTTGGAGGCCGGCGTGGGGAAGATG  240
                                                          M     1
241  GCCCAGTCCACCGCCACCTCCCCTGATGGGGGCACCACGTTTGAGCACCTCTGGAGCTCT  300
  2  A  Q  S  T  A  T  S  P  D  G  G  T  T  F  E  H  L  W  S  S    21
301  CTGGAACCAGACAGCACCTACTTCGACCTTCCCCAGTCAAGCCGGGGGAATAATGAGGTG  360
 22  L  E  P  D  S  T  Y  F  D  L  P  Q  S  S  R  G  N  N  E  V    41
361  GTGGGCGGAACGGATTCCAGCATGGACGTCTTCCACCTGGAGGGCATGActACATCTGTC  420
 42  V  G  G  T  D  S  S  M  D  V  F  H  L  E  G  M  T  T  S  V    61
421  ATGGCCCAGTTCAATCTGCTGAGCAGCACCATGGACCAGATGAGCAGCCGCGCGGCCTCG  480
 62  M  A  Q  F  N  L  L  S  S  T  M  D  Q  M  S  S  R  A  A  S    81
481  GCCAGCCCCTACACCCCAGAGCACGCCGCCAGCGTGCCCACCCACTCGCCCTACGCACAA  540
 82  A  S  P  Y  T  P  E  H  A  A  S  V  P  T  H  S  P  Y  A  Q   101
541  CCCAGCTCCACCTTCGACACCATGTCGCCGGCGCCTGTCATCCCCTCCAACACCGACTAC  600
102  P  S  S  T  F  D  T  M  S  P  A  P  V  I  P  S  N  T  D  Y   121
601  CCCGGACCCCACCACTTTGAGGTCACTTTCCAGCAGTCCAGCACGGCCAAGTCAGCCACC  660
122  P  G  P  H  H  F  E  V  T  F  Q  Q  S  S  T  A  K  S  A  T   141
661  TGGACGTA.....
142  W  T     .  .
```

FIG. 17

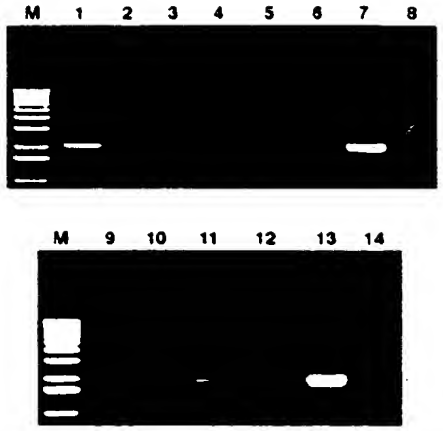

FIG. 18

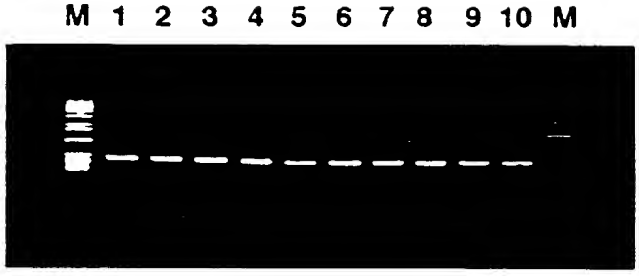

FIG. 19A

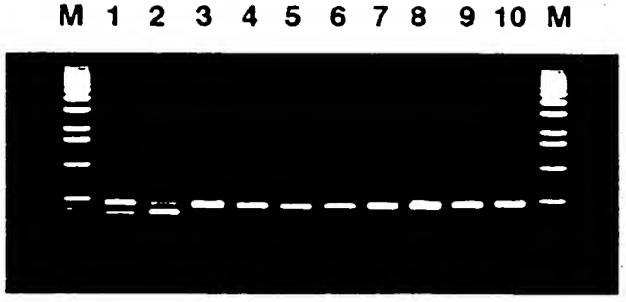

FIG. 19B

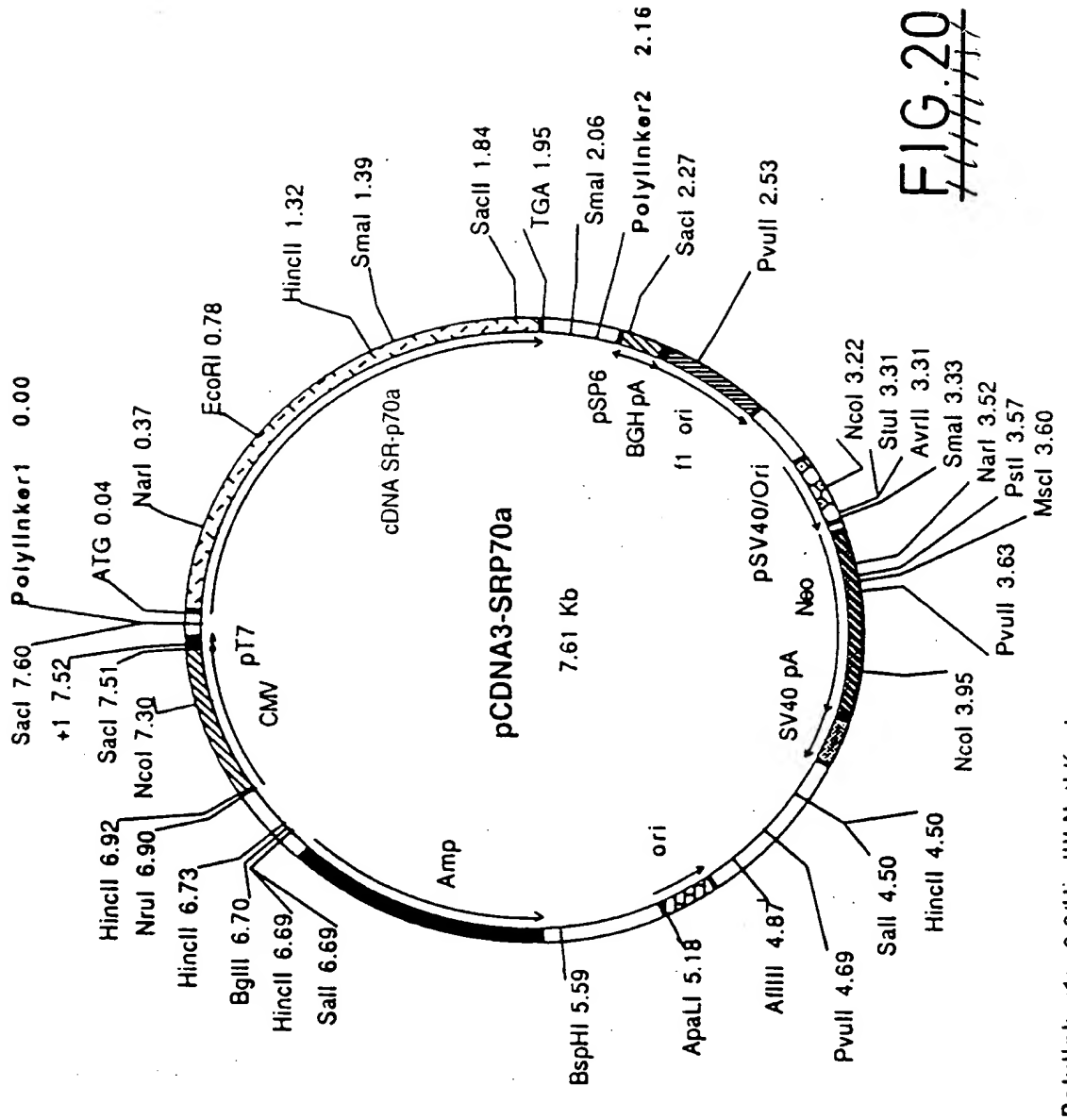

FIG.20

PolyIlnker1: 0.0/HindIII.NotI.KpnI.
PolyIlnker2: 2.16/XbaI.NotI.ApaI.